

A geometric study of many body systems

Eldar Straume

Department of Mathematical Sciences,
Norwegian University of Science and Technology (NTNU)
N-7491 Trondheim, Norway
e-mail: eldars@math.ntnu.no

February 7, 2008

Abstract

An n -body system is a labelled collection of n point masses in a Euclidean space, and their congruence and internal symmetry properties involve a rich mathematical structure which is investigated in the framework of equivariant Riemannian geometry. Some basic concepts are n -configuration, configuration space, internal space, shape space, Jacobi transformation and weighted root system. The latter is a generalization of the root system of $SU(n)$, which provides a bookkeeping for expressing the mutual distances of the point masses in terms of the Jacobi vectors. Moreover, its application to the study of collinear central n -configurations yields a simple proof of Moulton's enumeration formula. A major topic is the study of matrix spaces representing the shape space of n -body configurations in Euclidean k -space, the structure of the m -universal shape space and its $O(m)$ -equivariant linear model. This also leads to those "orbital fibrations" where $SO(m)$ or $O(m)$ act on a sphere with a sphere as orbit space. A few of these examples are encountered in the literature, e.g. the special case $S^5/O(2) \approx S^4$ was analyzed independently by Arnold, Kuiper and Massey in the 1970's.

Contents

1	Introduction	2
1.1	A brief overview	3
1.2	n -configuration space, internal space and shape space	4
1.3	Geometrization, symmetry and reduction	5
1.4	Congruence and internal symmetry in n -body spaces	7
1.5	Jacobi vectors and the centered configuration space	9
2	Jacobi transformations from a categorical viewpoint	11
2.1	Transformations of n -body systems and Jacobi's approach	11
2.2	Equivalent characterizations of Jacobi transformations	13
2.3	Construction of Jacobi vectors	15

3	Orthogonal transformation groups on matrix spaces	18
3.1	Compact transformation groups and orbital decomposition . . .	19
3.1.1	Basic definitions and constructions	19
3.1.2	Compact linear groups on Euclidean spaces	21
3.1.3	The configuration space for m-body systems in d-space .	22
3.2	Transformations of the space of symmetric matrices	23
3.3	Algebraic realization of orbit spaces and orbital stratification . .	25
3.3.1	Cross sections and canonical forms	26
3.3.2	Orthogonality of cross sections	27
3.3.3	Rank and subrank stratification of matrix spaces	27
3.4	Topology of the shape space and related spaces	30
3.4.1	Local and global topology	31
3.4.2	Branched coverings and generalized Hopf fibrations	34
3.4.3	The m-universal linear model	37
3.4.4	The simplest shape spaces	40
4	Geometric invariants of n-body systems	42
5	The weighted root system of an n-body system	46
5.1	Distance functions and the Ψ -root system	47
5.2	The standard weighted root system	48
5.3	Weighted root systems and their metric invariants	49
5.4	On the role of the mass distribution	50
6	Collinear central configurations revisited	52

1 Introduction

An *n-body system* is defined to be a labelled collection of n point masses (or particles) P_i of mass m_i in Euclidean 3-space \mathbb{R}^3 , and it is of general importance to find appropriate mathematical models to describe and analyze such a system. We have in mind few-body systems as well as many-body systems, ranging from different areas such as celestial and quantum mechanics or quantum chemistry (n-body problem, n-atomic molecules etc.). Despite the diversity of the applications they all share a fundamental underlying mathematical structure, in terms of kinematic concepts and internal space geometry, and the role of the mass distribution. Here we shall focus attention on these basic structures, in a modern geometric and topological setting with orthogonal transformation groups in the forefront. This approach also establishes similar results for point masses in any Euclidean spaces $\mathbb{R}^d, d \geq 2$.

In this introductory section we first give an overview of the paper, which take up several different topics. Then, in the following subsections we introduce some basic concepts and constructions which we shall return to later. Here the presentation is rather informal or expository, with some comments on the history.

1.1 A brief overview

Our approach is to combine symmetry and kinematic geometric principles in the framework of equivariant geometry, exhibiting the importance of the classical orthogonal transformation groups and the associated orbit space constructions. This enables us to investigate in a unifying way the notions of congruence, internal configuration space and internal symmetry group. Internal symmetries of n-body systems are investigated in Section 2 from a categorical viewpoint based on pure geometrical principles. As we shall explain, for various reasons they should be referred to as Jacobi transformations, but in the physics literature they are also encountered as *kinematic rotations* or *democracy transformations*, cf. [17].

Section 3 is largely devoted to the study of the topology of shape spaces in general, and here it is natural to consider n-body systems in higher dimensional Euclidean spaces \mathbb{R}^d as well. We shall exploit the fact that congruence and internal symmetry for point masses in \mathbb{R}^d combine together to a tensor product representation of some $O(d) \times O(m)$, acting on the matrix space $M(d, m) \simeq \mathbb{R}^{dm}$ by matrix multiplication. In short, a typical shape space is the orbit space $M^* = S^p/O(d)$, where S^p is the unit sphere of $M(d, m)$. Now M^* inherits the action of $O(m)$ as its symmetry group, but on the other hand, M^* also naturally embeds as an $O(m)$ -invariant subset of the linear space $Sym^0(m)$ of symmetric matrices with zero trace. Of particular interest are the cases where $S^p/H \approx S^q$ is also a sphere, $H = SO(m)$ or $O(m)$. They yield an infinite family of "orbital fibrations" $S^p \rightarrow S^q$, which for $q = 1, 2, 4$ are the Hopf fibrations, cf. Section 3.4.2.

In the study of many body systems, we apply the orbit space reduction to the configuration manifold rather than to its cotangent bundle, as in the reduction method of Marsden-Weinstein which, for example, Iwai [13] applies to the Hamiltonian system describing classical molecular dynamics. On the other hand, R. Littlejohn and his collaborators (cf. e.g. [16], [17], [18]) have investigated the gauge fields that arise on the reduced (i.e. internal) configuration space, and our work is directly related to the geometric framework of their investigations.

In Section 4 we determine the geometric invariants of n-body systems, namely polynomial functions on the configuration space which are both congruence invariants and internal symmetry (or democracy) invariants. The topic is certainly well understood in classical invariant theory, but we are also seeking symmetrized expressions for centered n-body configurations, that is, their center of mass is fixed at the origin.

In Section 5 we introduce the notion of the *weighted root system* of an n-body system with a given mass distribution. This is a geometric invariant which generalizes the notion of a root system of Cartan type A_{n-1} in classical Lie theory, and its underlying structure is actually inherent in various contexts. For example, it encodes the data of the relative positions of the binary collision varieties in the configuration space or shape space. We shall introduce it as a bookkeeping device for expressing the *mutual distances* between the n bodies in terms of the Jacobi vectors.

In Section 6 we present, as a nice application of the weighted root system, a simple proof of Moulton's classification of collinear *central n-configurations* (cf. [23]) which is also well adapted for numerical computations. Recall that these are the configurations characterizing n collinear masses capable of a rigid uniform rotation under the mutual gravitational forces.

The present paper is essentially a preprint (with the same title) of the author from April 2002, prompted by the paper Hsiang [9] and the succeeding joint work [11]. Our Sections 5-6 recall the contents of [11], but in the present paper the topology of general shape spaces, the universal shape space and its linear model, is a major topic. It turns out that some few-body shape spaces have an interesting history in the literature. For example, it was surprising to find that the quotient space of complex projective plane \mathbb{CP}^2 modulo complex conjugation is topologically a 4-sphere, namely $S^5/O(2) \approx S^4$. Several independent and different proofs of this fact had already been published (cf. Arnold [2], [3], Kuiper[15], Massey[21]). Now we also find that our Section 3 has some overlapping with the more recent paper Atiyah-Berndt [1].

1.2 n-configuration space, internal space and shape space

The location of an n-body system is conveniently represented by its *n-configuration* $\mathbf{X} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ where \mathbf{a}_i is the position vector of P_i . The n-tuple \mathbf{X} is regarded as a vector in the *free n-configuration space*

$$\hat{M}_n = \mathbb{R}^3 \times \mathbb{R}^3 \times \dots \times \mathbb{R}^3 \simeq \mathbb{R}^{3n}, \quad (1)$$

namely a Euclidean 3n-space with an orthogonal splitting reflecting the individual positions of the n point masses, which may possibly coincide. The mass distribution (m_1, m_2, \dots, m_n) is tacitly assumed to be fixed unless otherwise stated, and in general we assume $m_i > 0$. The *centered n-configuration space* is the subspace of dimension $3n - 3$

$$M_n \subset \hat{M}_n : \sum m_i \mathbf{a}_i = 0, \quad (2)$$

consisting of those n-configurations with its center of mass at the origin. \hat{M}_n has the following mass dependent (Jacobi) *kinematic metric*, namely the inner product of \mathbf{X} and $\mathbf{Y} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ is

$$\langle \mathbf{X}, \mathbf{Y} \rangle = \sum m_i \mathbf{a}_i \cdot \mathbf{b}_i \quad (3)$$

The isometry group of \hat{M}_n is the associated Euclidean group

$$E(\hat{M}_n) = \hat{M}_n \hat{\times} O(\hat{M}_n) \simeq \mathbb{R}^{3n} \hat{\times} O(3n) = E(3n) \quad (4)$$

which we have expressed in the usual way as a semidirect product of the subgroups of translations and orthogonal transformations respectively.

Two n-configurations \mathbf{X} and \mathbf{Y} are regarded as *congruent* if their i-th components \mathbf{a}_i and \mathbf{b}_i differ by the same Euclidean motion in \mathbb{R}^3 for each i. Thus the congruence relation is defined by the natural (diagonal) $E(3)$ -action on \hat{M}_n ,

by which $E(3)$ (resp. $O(3)$) is embedded as a subgroup of $E(3n)$ (resp. $O(3n)$). The space \bar{M}_n of congruence classes is the (*congruence*) *moduli space*, also referred to as the *internal space*. It may be viewed as the result of the two-step orbit space construction

$$\hat{M}_n \rightarrow \frac{\hat{M}_n}{\mathbb{R}^3} \simeq M_n \rightarrow \frac{M_n}{O(3)} = \frac{\hat{M}_n}{E(3)} = \bar{M}_n \quad (5)$$

where in the first step the translation-reduced space is identified with M_n . The second step is the orbit space construction of the transformation group $(O(3), M_n)$, from which it follows that \bar{M}_n is a stratified manifold of dimension $3n - 6$ (cf. Section 3).

Next, let us divide "congruence" into "shape" and "size" and use the squared norm function I (cf. (7)), namely the polar moment of inertia, as a measure of the size of an n -configuration (or of its congruence class). Then \bar{M}_n has the structure of a cone, where each ray emanates from the cone vertex (or base point) $O = (I = 0)$ and the ray represents a fixed shape. The *shape space* is the space of rays, which we may regard as the orbit space

$$\frac{\bar{M}_n - \{0\}}{\mathbb{R}^+} = \frac{M_n - \{0\}}{O(3) \times \mathbb{R}^+} \simeq M_n^* \quad (6)$$

where $O(3) \times \mathbb{R}^+$ is the group of similarity transformations of Euclidean 3-space with the induced (diagonal) action on M_n . However, each ray has a unique point where $I = 1$, and therefore it is more convenient to identify M_n^* with the subset of \bar{M}_n consisting of classes of unit size $I = 1$. Thus, the internal space \bar{M}_n is naturally a cone over the shape space M_n^* .

1.3 Geometrization, symmetry and reduction

The viewpoint that kinematics is a geometric discipline has a long history. For an n -body motion $t \rightarrow \mathbf{X}(t)$, the fundamental kinematic quantities are

$$\begin{aligned} I &= |\mathbf{X}|^2 = \sum m_i |\mathbf{a}_i|^2 \\ T &= \frac{1}{2} |\dot{\mathbf{X}}|^2 = \frac{1}{2} \sum m_i |\dot{\mathbf{a}}_i|^2 \\ \boldsymbol{\Omega} &= \mathbf{X} \times \dot{\mathbf{X}} = \sum m_i \mathbf{a}_i \times \dot{\mathbf{a}}_i \\ \mathbf{p} &= \sum m_i \dot{\mathbf{a}}_i \end{aligned} \quad (7)$$

These are the moment of inertia, kinetic energy, angular and linear momentum, respectively. In particular, the *hyperradius* $\rho = \sqrt{I}$ is the norm of \mathbf{X} in the n -configuration space (1) with respect to the inner product (3), and using T the same metric may be presented as the *kinematic Riemannian metric*

$$ds^2 = 2T dt^2 = \sum m_i (dx_i^2 + dy_i^2 + dz_i^2) \quad (8)$$

Furthermore, dynamics was incorporated in this geometric setting by the classical geometrization procedure which dates at least back to the early 19th

century. Let us briefly recall the basic idea behind this, namely Jacobi's reformulation of Lagrange's least action principle, which goes as follows. When dynamics is taken into account and the above n-body motion is due to a force field ∇U derived from the potential energy $-U$, the total energy $h = T - U$ is conserved and the metric (8) should be conformally modified to the following *dynamical metric*

$$ds_h^2 = (U + h)ds^2 \quad (9)$$

depending on a fixed energy level h . Then the trajectories of total energy h can be recovered as the geodesics of this metric. This applies, for example, to physically important n-body systems with potential functions (such as the Newtonian or Coulomb potential) depending on the pairwise distances $r_{ij} = |\mathbf{a}_i - \mathbf{a}_j|$. Therefore, they are invariant under the Euclidean group $E(3)$, and consequently both linear and angular momentum of the motion are conserved.

Since the early days it has been an important issue how to fully utilize conservation laws. Sometimes these are first integrals associated with symmetry groups, and the reduction of integration problems using continuous (or infinitesimal) symmetries dates back to Sophus Lie's work in the 1870's.

Clearly, the most effective usage of the invariance of linear momentum \mathbf{p} is to choose an inertial frame at the center of mass, thereby reducing the configuration space \hat{M}_n to its subspace M_n (2) and hence the associated linear momentum vanishes. A much harder problem is to further use the invariance of $\mathbf{\Omega}$, which is related to the congruence action of $O(3)$ (or $SO(3)$) on M_n and its cotangent bundle. Namely, if T^* denotes the cotangent bundle construction, consider the two ways of reducing the phase space

$$T^*\left(\frac{M_n}{O(3)}\right) = T^*\bar{M}_n, \quad \frac{T^*(M_n)}{O(3)}$$

to dimension $6n - 12$ and $6n - 9$, respectively. This suggests that the orbit space reduction is most effective at the level of the configuration space rather than the phase space.

In this paper we shall focus attention on two main issues. Firstly, it is important to construct appropriate coordinates for M_n , which is actually a problem with no canonical or generally "best" solution. However, from the representation theory of $O(3)$ and natural guidelines suggested by the splitting or invariance of the kinematic quantities, we are led to a natural approach whose origin may, in fact, be ascribed to Jacobi. Secondly, we inquire about the structure of the internal space \bar{M}_n , as an orbit space of M_n modulo $O(3)$. It has the natural *kinematic metric*

$$d\bar{s}^2 = 2\bar{T}dt^2 \quad (10)$$

where $\bar{T} = T - T^\omega$ and T^ω is the kinetic energy due to purely rotational motion. Indeed, this metric is Riemannian and coincides with the induced orbital distance metric (see e.g. [30]), and the reduction map $M_n \rightarrow \bar{M}_n$ is a (stratified) Riemannian submersion. As a consequence, the geodesics in \bar{M}_n are the image curves of those geodesics (i.e., linear motions of constant speed) in (M_n, ds^2) with vanishing angular momentum.

On the other hand, the appropriate reduction of the dynamical equations in M_n to the level of \bar{M}_n is generally a hard problem even today. In M_n the geometrization procedure yields the conformal modification of the kinematic metric (8) leading to the dynamical metric (9). Similarly, one may search for a similar geometrization procedure at the internal space level, which should yield a dynamical Riemannian metric on \bar{M}_n of type

$$d\bar{s}_{h,\Omega}^2 = F(U, h, \Omega) d\bar{s}^2 \quad (11)$$

in analogy with (9). However, this only works in special cases, for example, for the (classical) planar n-body problem, but we leave this topic here.

Actually, in his study of celestial mechanics Jacobi himself abandoned the above Riemannian geometric approach in favor of the increasingly successful Hamiltonian formalism where, at the phase space level (symplectic geometry), ideas involving symmetry, conservation laws and integrability questions have been continuously developed up to present time. On the other hand, with the modern techniques of equivariant geometry, Lie transformation groups and related reduction theory and quotient constructions, the framework of Riemannian geometry and modern differential geometric techniques are now more readily applicable for the study of n-body problems. We also refer to Littlejohn-Reinsch[17] for a general discussion with many references to the gauge kinematic and dynamics of many particle systems. For real historical background information on the geometrization of physics we also propose the survey article Lützen[19] on the 19th century interactions between mechanics and differential geometry.

1.4 Congruence and internal symmetry in n-body spaces

Let us explain the interaction of the notions of congruence and internal symmetry for n-body systems, where by congruence we mean $O(3)$ -congruence and assume the translational degrees of freedom have been eliminated. For this purpose we introduce a notion somewhat more general than a free m-configuration space (1), namely an *m-body space*

$$M \simeq \mathbb{R}^{3m} : (O(3), m\rho_3)$$

is a Euclidean 3m-dimensional space with an orthogonal representation of type $(O(3), m\rho_3)$, that is, m copies of the standard representation ρ_3 of $O(3)$. However, particles and mass distributions are not mentioned, and there is no specific decomposition into 3-dimensional invariant subspaces \mathbb{R}^3 as in (1). As an example, we have in mind the centered n-configuration space M_n (2), where $m = n - 1$ and the original mass distribution is disguised in the Euclidean metric (3).

On the other hand, associated with \mathbb{R}^{3m} is the totality $\mathcal{D}(\mathbb{R}^{3m})$ of all orthogonal, $O(3)$ -invariant and ordered decompositions

$$\mathbb{R}^{3m} = V_1 \oplus V_2 \oplus \dots \oplus V_m, \quad V_i \simeq \mathbb{R}^3. \quad (12)$$

Then, for a specific choice of decomposition we can identify \mathbb{R}^{3m} with the free m-configuration space \hat{M}_m as in (1) and associate a particle P_i with position

vector \mathbf{a}_i to each summand V_i . Moreover, the metric on \mathbb{R}^{3m} viewed as a metric on \hat{M}_m should have the form (3), so we also need a mass distribution. However, by suitably scaling of the vectors \mathbf{a}_i we are actually free to choose any mass distribution since a vector in \mathbb{R}^{3m} written as an m-configuration $(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m)$ determines position vector \mathbf{a}_i and mass m_i modulo the constraint $\mathbf{b}_i = \sqrt{m_i} \mathbf{a}_i$ for each i (see Section 2.1).

In any case, \mathbb{R}^{3m} has the given *congruence group* $O(3) \subset O(3m)$, and we define the *internal* (or inner) symmetries to be those transformations $\varphi \in O(3m)$ which commute with $O(3)$, that is, $\varphi\psi = \psi\varphi$ for each $\psi \in O(3)$. By standard representation theory (cf. Schur's lemma) these φ constitute a subgroup isomorphic to $O(m)$, which we shall refer to as the (internal) *symmetry group*. The two subgroups of $O(3m)$ intersect at $\mathbb{Z}_2 = \{\pm Id\}$ and hence combine to a subgroup

$$\frac{O(3) \times O(m)}{\mathbb{Z}_2} \rightarrow O(3m)$$

which is also described mathematically by the tensor product $\rho_3 \otimes \rho_m$ of their standard representations.

We also point out that $O(m)$ acts naturally on the set of decompositions (12). This action is, in fact, transitive and has isotropy group of type $O(1)^m$, and hence establishes a 1-1-correspondence

$$\mathcal{D}(\mathbb{R}^{3m}) \simeq \frac{O(m)}{O(1)^m}$$

To be more explicit, let us fix some decomposition (12) and choose orthonormal bases to identify each V_i with \mathbb{R}^3 , and hence identify \mathbb{R}^{3m} in (12) with the space $M(3, m)$ of real $3 \times m$ - matrices

$$X = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m) \quad (13)$$

with the standard Euclidean norm square

$$|X|^2 = \text{trace}(X^t X) = \sum_{i=1}^m |\mathbf{x}_i|^2, \quad (14)$$

where $\mathbf{x}_i \in V_i$ is the i-th column vector of X . Then the action of $\psi \in O(3)$ and $\varphi \in O(m)$ is just matrix multiplication on the left and right side respectively, inducing a joint left action on matrices by

$$(\psi, \varphi)X = \psi X \varphi^{-1} \text{ (matrix multiplication)} \quad (15)$$

Thus congruence and symmetry combined together is the following tensor product representation

$$(O(3) \times O(m), \rho_3 \otimes \rho_m, M(3, m)), \quad (16)$$

whose orbit structure will be analyzed by combining two consecutive orbit space constructions

$$M(3, m) \rightarrow \frac{M(3, m)}{O(3)} = \bar{M}(3, m) \rightarrow \frac{\bar{M}(3, m)}{O(m)} = \frac{M(3, m)}{O(3) \times O(m)} \simeq C(\Delta) \quad (17)$$

Of primary interest is the first orbit space, namely the *internal space* $\bar{M} = \bar{M}(3, m)$ with its induced metric structure. The symmetry group $O(m)$ descends faithfully to an induced transformation group on \bar{M} which is, in fact, its isometry group. Thus, the last step in (17) yields the final orbit space which is geometrically the Riemannian cone over a spherical triangle Δ (resp. a circular arc if $m = 2$) :

$$C(\Delta) = \begin{cases} \mathbb{R}^3/B_3 & \text{if } m \geq 3 \\ \mathbb{R}^2/B_2 & \text{if } m = 2 \end{cases} \quad (18)$$

Of course, this is a Euclidean cone. However, as indicated it is also the fundamental domain of the Weyl group B_3 (resp. B_2 if $m = 2$) in classical Lie theory, and this tells us that $C(\Delta)$ embeds isometrically into $M(3, m)$ as a cross section of the transformation group (16). Namely, $C(\Delta)$ hits every orbit of that group at a unique point and, moreover, it is perpendicular to every orbit.

We refer to Section 3 for further analysis of the above orbit spaces, in a more general setting involving all matrix spaces $M(d, m)$ with the transformation groups $O(d) \times O(m)$, for any $d \geq 1$.

1.5 Jacobi vectors and the centered configuration space

In dynamics it is the conservation of linear momentum \mathbf{p} that enables one to reduce the n-body problem to an (n-1)-body problem plus a trivial 1-body problem for the motion of the center of mass. To explain this, consider the canonical orthogonal and $O(3)$ -invariant decomposition of the free n-configuration space

$$\hat{M}_n = M_n \oplus \Delta\mathbb{R}^3, \text{ cf. (1)} \quad (19)$$

where the subspace $M_n \simeq \mathbb{R}^{3n-3}$ is the *centered* n-configuration space (2) and the "diagonal" $\Delta\mathbb{R}^3 = \{(\mathbf{a}, \mathbf{a}, \dots, \mathbf{a})\}$ is its orthogonal complement. For a motion $\mathbf{X}(t)$ in \hat{M}_n the vector $\mathbf{a} = \mathbf{a}(t)$ is, indeed, the center of mass of $\mathbf{X}(t)$. Now, conservation of \mathbf{p} means our inertial frame of reference will remain inertial if we translate the frame so that \mathbf{a} becomes the new origin. Therefore, with respect to the new frame, the motion will take place in the summand M_n in (19). This simple reduction is a key step in the integration of the classical Kepler problem, where $n = 2$ and $M_2 \simeq \mathbb{R}^3 \simeq \hat{M}_1$ is the configuration space of a fictitious 1-body system.

However, for $n > 2$ M_n is just an (n-1)-body space and there is no canonical way of further decomposing M_n to become the configuration space of $n - 1$ fictitious particles with appropriate position vectors and mass distribution, cf. (12). This fact is clearly reflected by the variety of types of coordinates for M_n which can be found in the literature. Recall, for example, the efforts of Lagrange, Jacobi and Delaunay who constructed their own "good" coordinates to study the classical 3-body problem (cf. e.g. Marchal [20]).

Of particular interest to us is Jacobi's approach, which by repeated applications generalizes to n bodies, but again there is no canonical way of doing so. Anyhow, our interpretation of his basic idea is that a solution of the above

splitting problem amounts to the construction of a transformation

$$\begin{aligned} \Psi : M_n &\rightarrow \hat{M}_{n-1} \\ &: (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) \rightarrow (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1}) \end{aligned} \quad (20)$$

with the "appropriate" properties (see below), connecting the $(n-1)$ -body space M_n (with the Jacobi metric (3)) to a *standard model*, namely the free configuration space \hat{M}_{n-1} with all masses equal to 1.

By viewing the vectors $\mathbf{x}_i \in \mathbb{R}^3$ as the columns of a matrix (13) we may identify \hat{M}_{n-1} with the matrix space

$$M(3, n-1) = \mathbb{R}^3 \otimes \mathbb{R}^{n-1} = \sum_{i=1}^{n-1} \oplus \mathbb{R}_i^3 = \sum_{j=1}^3 \oplus \mathbb{R}_j^{n-1} \quad (21)$$

with the norm as in (14), congruence group $O(3)$, symmetry group $O(n-1)$ and their joint tensor product representation of $O(3) \times O(n-1)$, see (15), (16). In (21) we have also indicated that the column and row vectors of a matrix belong to two different Euclidean spaces, having the standard action of $O(3)$ and $O(n-1)$ respectively.

Now, what are those "appropriate" properties Jacobi transformations such as Ψ should satisfy? In short, the answer is that Ψ is just an $O(3)$ -equivariant isometry, that is,

$$i) \Psi(g\mathbf{X}) = g\Psi(\mathbf{X}), \quad g \in O(3), \quad ii) |\mathbf{X}| = |\Psi(\mathbf{X})| \quad (22)$$

Let Ψ in (20) be a given transformation of this type. It associates to a centered n -configuration $\mathbf{X} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ its *Jacobi vector matrix*

$$X = \Psi(\mathbf{X}) = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1}) \in M(3, n-1)$$

whose column vectors \mathbf{x}_i will be referred to as the corresponding *Jacobi vectors*. On the other hand, if Ψ' is another Jacobi transformation, then the composition $\Psi' \circ \Psi^{-1} = \varphi$ is still an $O(3)$ -equivariant isometry. Consequently, we can write $\Psi' = \varphi \circ \Psi$ where

$$\varphi : M(3, n-1) \rightarrow M(3, n-1) \quad (23)$$

is a Jacobi transformation of the standard model (21), namely an orthogonal transformation commuting with $O(3)$. In other words, φ belongs to the symmetry group $O(n-1)$ (acting on the matrices X by right multiplication). This explains the non-uniqueness of Ψ ! Briefly, by knowing only one of the Ψ 's we obtain all of them by composing Ψ with any $\varphi \in O(n-1)$.

So far, however, we have not constructed a single Ψ in (20), but we have just seen that this is all we need to do. We refer to Section 2.2 for the explicit construction of our *standard* choice $\Psi = \Psi_0$ of Jacobi transformation and hence also our preferred choice of Jacobi vectors \mathbf{x}_i as distinguished linear combinations of the vectors \mathbf{a}_j .

Remark 1.1 *Quantities which are invariant under the transformation group $(O(n-1), M_n)$ are also insensitive to the different orderings of the n bodies. Therefore, this group is sometimes referred to as the democracy group. Elements $\varphi \in SO(n-1)$ are also named kinematic rotations. See e.g. [17].*

2 Jacobi transformations from a categorical viewpoint

Loosely speaking, there is the category of m -body spaces whose objects are the Euclidean spaces $M \simeq \mathbb{R}^{3m}$ with an orthogonal transformation group of type $(O(3), m\rho_3)$, and the morphisms will be called *Jacobi transformations*, namely they are the $O(3)$ -equivariant isometries $M \rightarrow M'$ between the spaces. Here the matrix space $M(3, m) = \hat{M}_m$ is a distinguished object which also serves as our *standard model* (21). Therefore, for a given m -body space M a "good" coordinate system for M amounts to the choice of an appropriate Jacobi transformation

$$\Psi : M \rightarrow M(3, m)$$

and the columns of the matrix $X \in M(3, m)$ will be the Jacobi vectors with respect to Ψ . We are primarily interested in the centered n -configuration space $M = M_n$ (where $m = n - 1$).

In the following two subsections we shall characterize *Jacobi transformations* in several equivalent ways, and we start with the classical Jacobi construction which also justifies our usage of terminology.

2.1 Transformations of n -body systems and Jacobi's approach

In the Introduction we actually started (less formally) with the category of n -body systems, whose objects are n -tuples (P_1, P_2, \dots, P_n) of point masses $P_i = (\mathbf{a}_i, m_i)$ with \mathbf{a}_i and m_i as position vector and mass respectively. In this setting, what should be the appropriate morphisms

$$\Phi : (P_1, \dots, P_n) \rightarrow (P'_1, \dots, P'_n) \quad ? \quad (24)$$

After all, the usefulness of such transformations with the desired properties, to simplify the further analysis, is well documented in both classical and quantum mechanics. First of all, we propose to consider transformations

$$\Phi : \begin{array}{l} (\mathbf{a}_1, \dots, \mathbf{a}_n) \rightarrow (\mathbf{a}'_1, \dots, \mathbf{a}'_n) \\ (m_1, \dots, m_n) \rightarrow (m'_1, \dots, m'_n) \end{array} \quad (25)$$

where the first map in (25) is an invertible linear transformation on $\hat{M}_n = \mathbb{R}^{3n}$ which may possibly depend on the masses m_i , with the choice of "new" masses m'_i constrained in some way. But Φ also transforms the motions of n -body system and hence also the first three basic kinematic quantities in (7), so let us demand that Φ preserves them, namely

$$\begin{aligned} I &= \sum m_i |\mathbf{a}_i|^2 = \sum m'_j |\mathbf{a}'_j|^2 \\ 2T &= \sum m_i |\dot{\mathbf{a}}_i|^2 = \sum m'_j |\dot{\mathbf{a}}'_j|^2 \\ \Omega &= \sum m_i \mathbf{a}_i \times \dot{\mathbf{a}}_i = \sum m'_j \mathbf{a}'_j \times \dot{\mathbf{a}}'_j \end{aligned} \quad (26)$$

Motivated by an idea of Jacobi, we define the (i, j) -*basic Jacobi transformation* for $i \neq j$

$$\rho_{ij} : (P_1, \dots, P_n) \rightarrow (P'_1, \dots, P'_n) \quad (27)$$

by demanding $P'_k = P_k, m'_k = m_k$ for $k \neq i, j$, and

$$\begin{aligned} P'_i &= (\mathbf{a}'_i, m'_i) = (\mathbf{a}_i - \mathbf{a}_j, \frac{m_i m_j}{m_i + m_j}) \\ P'_j &= (\mathbf{a}'_j, m'_j) = (\frac{m_i \mathbf{a}_i + m_j \mathbf{a}_j}{m_i + m_j}, m_i + m_j) \end{aligned} \quad (28)$$

In particular, P'_j is the center of mass of P_i and P_j , and m'_i is their reduced mass

$$\mu_{ij} = \frac{m_i m_j}{m_i + m_j} \quad (29)$$

Among the basic Jacobi transformations let us also include the following mass normalizing transformation

$$\rho_0 : (\mathbf{a}_i, m_i) \rightarrow (\sqrt{m_i} \mathbf{a}_i, 1), \forall i \quad (30)$$

and the permutation transformations

$$\bar{\sigma} : (P_1, \dots, P_n) \rightarrow (P_{\sigma(1)}, \dots, P_{\sigma(n)}) \quad (31)$$

which permute the vectors \mathbf{a}_i and masses m_i covariantly, namely there is one for each permutation σ of $\{1, 2, \dots, n\}$. It is straightforward to check that the above transformations, and hence all their compositions, have the invariance property (26).

Next, we inquire whether the above special transformations already generate all those Φ with the property (26). This is affirmatively settled in the following subsection, but first we shall make some clarifying observations. Note that velocity vectors $\dot{\mathbf{a}}_i$ of a motion are transformed in the same way as the position vectors \mathbf{a}_i and may well be regarded as vectors \mathbf{b}_i independent of the \mathbf{a}_i 's. In particular, invariance of I means the same as invariance of T , namely that Φ is an isometry. Therefore, the remaining issue is the nature of the invariance of Ω .

Consider the $3n \times 3n$ -matrix of a linear transformation Φ , regarded as an $n \times n$ -matrix

$$[\Phi] = \begin{pmatrix} B_{11} & B_{12} & \dots \\ B_{21} & B_{22} & \dots \\ \dots & \dots & \dots \end{pmatrix} \quad (32)$$

where each block B_{ij} is a 3×3 -matrix. Such an Φ expresses each \mathbf{a}'_j as a linear combination of the old ones, that is,

$$\mathbf{a}'_i = \sum_{j=1}^n \beta_{ij} \mathbf{a}_j, \quad i = 1, \dots, n, \quad (33)$$

if and only if each 3-block is a *scaling matrix* $B_{ij} = \beta_{ij} Id$, where Id denotes the identity matrix. Equivalently, Φ commutes with the (diagonal) action of $GL(3)$ on \mathbb{R}^{3n} and hence (by representation theory) belongs to the subgroup $GL(n)$ embedded into $GL(3n)$ by the tensor product action of $GL(3) \times GL(n)$.

On the other hand, assuming (33) holds Φ will be an isometry (i.e. I is preserved) if and only if the $n \times n$ -matrix (β_{ij}) is "orthogonal" in the sense that

$$(\beta_{ij}) \in D'^{-1}O(n)D \subset GL(n) \quad (34)$$

where $D = \text{diag}(\sqrt{m_1}, \dots, \sqrt{m_n})$, $D' = \text{diag}(\sqrt{m'_1}, \dots, \sqrt{m'_n})$. Finally, assuming (33) and (34) it is easy to verify that Ω is also preserved.

2.2 Equivalent characterizations of Jacobi transformations

It turns out that the invariance properties (26) are satisfied solely by demanding invariance of the vector Ω . This is elucidated by the following proposition where the angular momentum construction is analyzed from a purely algebraic viewpoint.

Proposition 2.1 *Consider two n -tuples of vectors in 3-space*

$$X = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n), Y = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n)$$

and define their angular momentum vector by

$$\Omega = \Omega(X, Y) = \sum_{i=1}^n \mathbf{x}_i \times \mathbf{y}_i$$

Regard X, Y as vectors in $\mathbb{R}^{3n} = \mathbb{R}^3 \times \dots \times \mathbb{R}^3$ and consider general linear transformations $\Phi : \mathbb{R}^{3n} \rightarrow \mathbb{R}^{3n}$,

$$X \rightarrow X' = (\mathbf{x}'_1, \mathbf{x}'_2, \dots, \mathbf{x}'_n), \quad Y \rightarrow Y' = (\mathbf{y}'_1, \mathbf{y}'_2, \dots, \mathbf{y}'_n)$$

Then Φ preserves angular momentum, that is, for all X, Y

$$\Omega(X, Y) = \Omega(X', Y')$$

if and only if Φ is an $O(3)$ -equivariant isometry.

Proof. The 3×3 -matrices

$$S_k = (s_{ij}^{(k)}), \quad k = 1, 2, 3,$$

whose only non-zero entries are $s_{ij}^{(k)} = -s_{ji}^{(k)} = 1$ whenever (i, j, k) is an even permutation of $(1, 2, 3)$, constitute a basis for the Lie algebra $so(3)$ of all skew-symmetric matrices S . The action of $SO(3)$ on \mathbb{R}^{3n} embeds each S as a "diagonal block" matrix

$$\Delta(S) \in so(3n)$$

with n copies of S along the diagonal. Observe that the matrices $\Delta(S_k)$ actually represent the three components Ω_k of the vector Ω , as skew-symmetric bilinear forms on \mathbb{R}^{3n} expressed as matrix products, namely

$$\Omega_k(X, Y) = X^t \Delta(S_k) Y$$

where X, Y are regarded as column vectors with $3n$ components.

Let us identify a transformation Φ with its matrix in $GL(3n)$. Those transformations leaving Ω_k invariant constitute a matrix group G_k defined by the constraint

$$\Phi^t \Delta(S_k) \Phi = \Delta(S_k)$$

and hence the group leaving Ω invariant is the intersection $G = G_1 \cap G_2 \cap G_3$ defined by the constraint

$$\Phi^t \Delta(S) \Phi = \Delta(S), \text{ for all } S \in so(3) \quad (35)$$

The proof amounts to show $G = O(n) \subset O(3n)$, namely that G is the group of isometries commuting with $O(3)$ (or equivalently with $SO(3)$).

If we knew Φ is orthogonal, then we could have applied the exponential function to (35) to conclude that Φ commutes with $SO(3)$ and hence belongs to $O(n)$. We claim, however, the isometry condition is itself a consequence of (35). To see this, write a typical matrix of the Lie algebra $L(G)$ as an $n \times n$ -matrix (A_{ij}) with blocks A_{ij} , see (32). Then the infinitesimal version of (35) simply reads

$$A_{ij}^t S + S A_{ij} = 0, \text{ for all } i, j \text{ and all } S \in so(3),$$

or equivalently

$$A_{ij} = -A_{ji}, \quad A_{ij} = \alpha_{ij} Id \quad (\text{scaling matrix})$$

However, these are precisely the conditions defining the Lie subalgebra $so(n)$ and consequently G and $O(n)$ have the same connected component $SO(n)$.

Finally, to show $G = O(n)$ we use the fact that G lies in the normalizer of $SO(n)$ in $GL(3n)$, namely in the image of $GL(3) \times O(n)$ by the tensor product representation. Therefore, we may assume $\Phi = \Delta(B)$ where $B \in GL(3)$ and hence by (35)

$$B^t S B = S \text{ for all } S \in so(3)$$

This condition says B leaves invariant all skew-symmetric bilinear forms in \mathbb{R}^3 , and consequently $B = \pm Id$. This completes the proof. ■

The basic Jacobi transformations (27), (30), (31) are maps between n -configuration spaces with possibly different mass distribution. However, using normalizing transformations ρ_0 , as in (30), any composition may be "pulled back" to a transformation $M(3, n) \rightarrow M(3, n)$ and with all masses $m_i = 1$. In this interpretation, the (i, j) -basic Jacobi transformation ρ_{ij} will be a (2-dimensional) rotation $\hat{\rho}_{ij} \in O(n)$ and it is easy to check that $\hat{\rho}_{12}, \hat{\rho}_{23}, \hat{\rho}_{13}$ generate a dense subgroup of $SO(3) \subset SO(3) \times SO(n-3) \subset O(n)$. Therefore, all rotations $\hat{\rho}_{ij}$ together with the permutations (31) generate a dense subgroup of $O(n)$.

As a consequence of the above observations we now state the following result on the characterization of *Jacobi transformations*, a terminology justified by (iii) below. In view of Section 1.3 we shall not attempt to define the angular momentum vector Ω in the broad category of n -body spaces, but in the following theorem we refer to n -configuration spaces as defined in Section 1.1.

Theorem 2.2 *The following four classes of linear transformations Φ between n -configuration spaces are identical :*

- (i) Φ is an isometry which preserves the angular momentum vector $\mathbf{\Omega}$.
- (ii) Φ preserves $\mathbf{\Omega}$.
- (iii) Φ is in the closure of the set of transformations generated by the basic Jacobi transformations.
- (iv) Φ is an $O(3)$ -equivariant isometry.

The Jacobi transformations on a fixed (n-body or n-configuration) space are also its symmetries (cf. Section 1.3), and they constitute a group isomorphic to $O(n)$. In particular, the symmetry group for our standard model, namely the matrix space $M(3, n)$, is the group $O(n)$ acting by matrix multiplication on the right side.

Transformations which preserve angular momentum and kinetic energy are certainly useful in quantum mechanics. Here one tries to keep operators "separable", that is, with no cross terms (cf. e.g. [6], §10.1). As a consequence of the above results Jacobi transformations provide all possible linear combinations \mathbf{a}'_j of the vectors \mathbf{a}_i which preserve the kinetic energy operator, in the sense that

$$T = -\frac{\hbar^2}{2} \left[\frac{1}{m_1} \nabla_{\mathbf{a}_1}^2 + \dots + \frac{1}{m_n} \nabla_{\mathbf{a}_n}^2 \right] = -\frac{\hbar^2}{2} \left[\frac{1}{m'_1} \nabla_{\mathbf{a}'_1}^2 + \dots + \frac{1}{m'_n} \nabla_{\mathbf{a}'_n}^2 \right] \quad (36)$$

A reduction of variables is achieved by choosing the transformation Φ so that \mathbf{a}'_n becomes the center of mass vector and hence vanishes relative to a center of mass coordinate system. Then, for example, the (time dependent) Schrödinger equation with potential energy V reduces to

$$i\hbar \frac{\partial}{\partial t} \Psi = \left[-\frac{\hbar^2}{2} \left(\sum_{i=1}^{n-1} \frac{1}{m'_i} \nabla_{\mathbf{a}'_i}^2 \right) + V(\mathbf{a}'_1, \dots, \mathbf{a}'_{n-1}) \right] \Psi$$

where $\Psi = \Psi(\mathbf{a}'_1, \dots, \mathbf{a}'_{n-1}, t)$. In the following subsection an explicit construction of such coordinates is given, and by usage of the symmetry group $O(n)$ they can be modified to satisfy additional properties for a specific purpose.

On the other hand, in the study of atomic structures one of the masses, say m_n (the nucleus), may be relatively large and then non-Jacobian coordinates may also turn out to be useful. However, in that case *mass polarization terms* cannot be avoided, namely mixed terms $\nabla_{\mathbf{a}'_i} \cdot \nabla_{\mathbf{a}'_j}$ will appear in the kinetic energy operator (see e.g. Appendix 8 in [5]).

2.3 Construction of Jacobi vectors

Here we shall construct an explicit Jacobi transformation (20), or rather its inverse

$$\Phi = \Psi^{-1} : X = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1}) \rightarrow (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{n-1}, -\frac{1}{m_n} \sum_{i=1}^{n-1} m_i \mathbf{a}_i) \quad (37)$$

from $M(3, n-1)$ to the centered configurations space M_n . By definition, Φ must be an $O(3)$ -equivariant isometry, and hence in analogy with (22)

$$\begin{aligned} i) \quad & \Phi(gX) = g\Phi(X), \quad g \in O(3) \\ ii) \quad & |\Phi(X)| = |X|, \quad X \in M(3, n-1) \end{aligned} \quad (38)$$

and let us also add the following "normalizing" condition :

$$(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k, 0, \dots, 0) \rightarrow (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k, \mathbf{b}^{(k)}, \dots, \mathbf{b}^{(k)}), \quad \text{for } k < n-1 \quad (39)$$

Namely, the Jacobi vectors \mathbf{x}_i should vanish for $i > k$ if and only if the particles P_i "collide" for $i > k$, that is, they occupy the same position $\mathbf{b}^{(k)}$. Equivalently, the Jacobi vectors define an orthogonal decomposition

$$M_n = \mathbb{R}_1^3 \oplus \mathbb{R}_2^3 \oplus \dots \oplus \mathbb{R}_{n-1}^3$$

where $\mathbb{R}_k^3 \simeq \mathbb{R}^3$ is the image of the imbedding

$$(0, \dots, 0, \mathbf{x}_k, 0, \dots, 0) \rightarrow (0, \dots, 0, \mathbf{a}_k - \mathbf{b}^{(k-1)}, \mathbf{b}_{(k)}, \dots, \mathbf{b}_{(k)}) \quad (40)$$

which actually represents a collinear 3-body system in the sense that P_1, P_2, \dots, P_{k-1} are located at the origin and P_{k+1}, \dots, P_n at the common position vector

$$\mathbf{b}_{(k)} = \mathbf{b}^{(k)} - \mathbf{b}^{(k-1)}.$$

The condition that $\mathbf{X} = (\mathbf{a}_1, \dots, \mathbf{a}_k, \mathbf{b}^{(k)}, \dots, \mathbf{b}^{(k)})$ in (39) belongs to M_n yields the formula

$$\mathbf{b}^{(k)} = \frac{-1}{m^{(k)}} \sum_{i=1}^k m_i \mathbf{a}_i, \quad \mathbf{b}^{(0)} = 0, \quad (41)$$

where by definition

$$m^{(k)} = m_{k+1} + \dots + m_n, \quad m^{(0)} = \sum m_i = \bar{m}$$

As a consequence of (39) we also note that the matrix $[\Phi]$ formally representing Φ by matrix multiplication

$$[\Phi] \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_{n-1} \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_{n-1} \end{pmatrix} \quad (42)$$

is lower triangular.

By combining (40) with the $O(3)$ -equivariance of Φ we obtain the following Jacobi vectors

$$\mathbf{x}_k = \zeta_k(\mathbf{a}_k - \mathbf{b}^{(k-1)}) = \zeta_k(\mathbf{a}_k + \frac{1}{m^{(k-1)}} \sum_{i=1}^{k-1} m_i \mathbf{a}_i) \quad (43)$$

for suitable constants $\zeta_k \neq 0$, whose square is determined by the isometry condition ii) of (38), namely

$$\zeta_k^2 = \frac{m_k m^{(k-1)}}{m^{(k)}}, \quad 1 \leq k \leq n-1 \quad (44)$$

For example, for $k=1$ we have $\mathbf{x}_1 = \zeta_1 \mathbf{a}_1$ with

$$|\mathbf{x}_1|^2 = \zeta_1^2 |\mathbf{a}_1|^2 = \left| (\mathbf{a}_1, \mathbf{b}^{(1)}, \dots, \mathbf{b}^{(1)}) \right|^2 = |\mathbf{a}_1|^2 \left(m_1 + \frac{m_1^2}{(\bar{m} - m_1)} \right)$$

and hence $\zeta_1^2 = m_1 \bar{m} / (\bar{m} - m_1)$.

Now it is easy to describe all Φ satisfying the condition (39), namely Φ is uniquely determined by the choice of sign of the numbers ζ_k , $1 \leq k \leq n-1$. There are 2^{n-1} such choices, and a natural choice is to take all $\zeta_k > 0$, namely our distinguished $\Phi = \Phi_0$ is defined by taking

$$\zeta_k = \sqrt{\frac{m_k m^{(k-1)}}{m^{(k)}}}, \quad 1 \leq k \leq n-1 \quad (45)$$

Its inverse $\Psi_0 = \Phi_0^{-1}$ will be regarded as our *standard* Jacobi transformation, and hence the *standard* Jacobi vectors are the columns of the matrix $X = \Psi_0(\mathbf{X})$. Then the corresponding lower triangular matrix $[\Psi_0] = [\Phi_0]^{-1}$ (see (42)), has entries

$$[\Psi_0]_{ik} = \zeta_i \cdot \begin{cases} 1 & i = k \\ \frac{m_k}{m^{(i-1)}}, & 1 \leq k < i \leq n-1 \\ 0 & 1 \leq i < k \leq n-1 \end{cases} \quad (46)$$

derived from the relation (43). On the other hand, the entries of $[\Phi_0] = (l_{ik})$ are

$$[\Phi_0]_{ik} = l_{ik} = \zeta_k^{-1} \cdot \begin{cases} 1 & i = k \\ -\frac{m_k}{m^{(k)}}, & 1 \leq k < i \leq n-1 \\ 0 & 1 \leq i < k \leq n-1 \end{cases} \quad (47)$$

where in each column the entries below the diagonal are identical. Moreover, the square sum for each row is given by the simple formula

$$l_{k1}^2 + \dots + l_{kk}^2 = \frac{1}{\bar{m}} \frac{\bar{m} - m_k}{m_k}, \quad 1 \leq k \leq n-1$$

Theorem 2.3 *The following Jacobi transformation*

$$\tilde{\Psi}_0 = \rho_0 \circ \rho_{1n} \circ \rho_{2n} \circ \dots \circ \rho_{n-1,n} : \hat{M}_n \rightarrow M(3, n)$$

maps $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ to $(\mathbf{x}_1, \dots, \mathbf{x}_n)$, where $\mathbf{x}_n = \bar{m}^{-1} \sum m_i \mathbf{a}_i$ is the center of mass and the remaining \mathbf{x}_i are given by the expressions (43), (45). In particular, the restriction of $\tilde{\Psi}_0$ to the centered configuration space yields $\mathbf{x}_n = 0$ and it coincides with the standard Jacobi transformation

$$\Psi_0 : M_n \rightarrow M(3, n-1) : (\mathbf{a}_1, \dots, \mathbf{a}_n) \rightarrow (\mathbf{x}_1, \dots, \mathbf{x}_{n-1})$$

The proof follows by an explicit calculation of \mathbf{x}_k as a linear combination of $\mathbf{a}_1, \dots, \mathbf{a}_k$, which shows that \mathbf{x}_k coincides with the vector in (43) for $k < n$, and furthermore, \mathbf{x}_n is the center of mass of the n -configuration $\mathbf{X} = (\mathbf{a}_1, \dots, \mathbf{a}_n)$.

The above transformation $\Psi_0 : M_n \rightarrow M(3, n-1)$ turns out, in fact, to be identical with the coordinate transformation described in Theorem 1 of Hsiang [9], derived by a similar but different "naturality" principle. For convenience, explicit formulae (43) for the *standard* Jacobi vectors when $n = 3, 4$ are listed below :

$$\begin{aligned}
n = 3 : \quad \mathbf{x}_1 &= \sqrt{\frac{m_1 \bar{m}}{m_2 + m_3}} \mathbf{a}_1, \quad \mathbf{x}_2 = \sqrt{\frac{m_2(m_2 + m_3)}{m_3}} \left(\mathbf{a}_2 + \frac{m_1}{m_2 + m_3} \mathbf{a}_1 \right) \\
n = 4 : \quad \mathbf{x}_1 &= \sqrt{\frac{m_1 \bar{m}}{m_2 + m_3 + m_4}} \mathbf{a}_1 \\
\mathbf{x}_2 &= \sqrt{\frac{m_2(m_2 + m_3 + m_4)}{m_3 + m_4}} \left(\mathbf{a}_2 + \frac{m_1}{m_2 + m_3 + m_4} \mathbf{a}_1 \right) \\
\mathbf{x}_3 &= \sqrt{\frac{m_3(m_3 + m_4)}{m_4}} \left[\mathbf{a}_3 + \frac{1}{m_3 + m_4} (m_1 \mathbf{a}_1 + m_2 \mathbf{a}_2) \right]
\end{aligned} \tag{48}$$

3 Orthogonal transformation groups on matrix spaces

By applying a fixed Jacobi transformation $\Psi : M_n \rightarrow M(3, n-1)$, quantities and constructions in the centered configuration space M_n are transported to its matrix model $M(3, m)$ with $m = n-1$, as explained in Section 1.4. In this setup we may and shall, however, discuss more generally the matrix space

$$M = M(d, m) \simeq \mathbb{R}^d \otimes \mathbb{R}^m$$

with its Euclidean norm (14) and orthogonal transformation group

$$G = G_1 \times G_2 = O(d) \times O(m)$$

acting by the tensor product $\rho_d \otimes \rho_m$, namely $(\psi, \varphi) \in G$ acts on matrices X by

$$((\psi, \varphi), X) \rightarrow \psi X \varphi^{-1} \text{ (matrix multiplication)} \tag{49}$$

Physically, we have in mind M as the configuration space for m free bodies (or $m+1$ bodies with fixed center of mass) in d -dimensional Euclidean space, and the orbit space $\bar{M} = M/G_1$ is the internal configuration space with the induced action of G_2 as the internal symmetries. Therefore, for further investigation of the topology of \bar{M} it is natural to analyze its induced G_2 -orbital structure. Thus we shall investigate two successive equivariant systems (G_1, M) and (G_2, \bar{M}) , where by an *equivariant system* (K, Y) we mean in general a (compact) Lie group K acting on a space Y . We also say (K, Y) is a *transformation group* or we simply say Y is a K -space, and it is convenient to refer to the orbit space construction and the orbit map

$$\pi : Y \rightarrow Y/K \tag{50}$$

as a *K-orbital fibration*. In (50) the "fibers" are the orbits, and therefore π may not be the projection of a fiber bundle (or fibration) in the usual sense since in general the orbit type is not unique (see below).

3.1 Compact transformation groups and orbital decomposition

First we shall recall some basic facts concerning equivariant systems (K, \mathcal{M}) where K is a compact Lie group acting smoothly on a manifold \mathcal{M} , that is, the action map $K \times \mathcal{M} \rightarrow \mathcal{M}$ is C^∞ -smooth. As a good reference on this topic we propose, for example, the book Bredon[7]. Then we specialize to the compact linear groups, and of particular interest are the natural transformation groups on the matrix spaces.

3.1.1 Basic definitions and constructions

The space \mathcal{M} splits into a disjoint union of K -orbits; these are minimal (or homogeneous) K -spaces in the obvious sense. The isotropy (or stability) groups K_p at points p along the same orbit constitute a single conjugacy class (K_α) of subgroups representing the *type* of the orbit. For simplicity we say the orbit is of type K_α , and as a K -space such an orbit is naturally equivalent to the coset space K/K_α . By grouping together orbits of the same type, \mathcal{M} decomposes as a union of orbit strata which are smooth submanifolds \mathcal{M}_α , all of which have dimension less than \mathcal{M} except the *principal* stratum \mathcal{M}_ω , which is an open and dense submanifold. The corresponding principal type K_ω is characterized as the unique smallest in the sense that (up to conjugation) $K_\omega \subset K_\alpha$ for each α .

The orbit map

$$\pi : \mathcal{M} \rightarrow \bar{\mathcal{M}} = \mathcal{M}/K \quad (51)$$

induces a smooth functional structure on the orbit space, in the sense that a function f on $\bar{\mathcal{M}}$ is called *smooth* if the composition $f \circ \pi$ is a smooth function on \mathcal{M} . Then $\bar{\mathcal{M}}$ is a *differential space*, but it may not be even locally Euclidean and hence it is not a manifold in general. However, still it has the nice and rich structure of a *stratified smooth manifold* with the image sets $\bar{\mathcal{M}}_\alpha = \pi(\mathcal{M}_\alpha)$ as strata, and they are actually smooth manifolds. In this way we may regard \mathcal{M} as a finite union of smooth *orbit bundles* denoted by

$$K/K_\alpha \hookrightarrow \mathcal{M}_\alpha \xrightarrow{\pi_\alpha} \bar{\mathcal{M}}_\alpha \quad (52)$$

whose fibers are those orbits of a fixed type K_α and π_α is, in fact, the projection of an actual fiber bundle.

For convenience, let us recall the general notion of a (locally trivial) fiber bundle, typically illustrated by a sequence like (52)

$$F \hookrightarrow E \xrightarrow{\pi} B \quad (53)$$

where π is the projection map, E (resp. B) is the total (resp. base) space, and F denotes the typical fiber, that is, the fibers $E_b = \pi^{-1}(b)$ are homeomorphic to F for each $b \in B$. The simplest example is the product bundle, where $E = B \times F$ and π is the obvious projection. An isomorphism between two bundles over B

is given by a fiber preserving homeomorphism $\varphi : E_1 \rightarrow E_2$ between their total spaces, and a bundle is *trivial* if it is isomorphic to a product bundle. The local triviality property (which is part of the definition) means that B is covered by open sets U such that the portion $\pi^{-1}(U)$ over U is a trivial bundle, that is, $\pi^{-1}(U) \simeq U \times F$. Finally, the bundle is smooth if all the above spaces are smooth manifolds and the maps involved are smooth.

In particular, a K -orbit bundle like (52) is defined by an equivariant system (K, E) with a single orbit type, and moreover, B is the orbit space E/K and hence the fibers are the K -orbits. A *principal bundle* (or principal fibration) is the special case

$$K \hookrightarrow E \rightarrow E/K = B \quad (54)$$

of (52) where K acts freely (i.e. all isotropy groups are trivial).

When K acts by isometries on a Riemannian manifold \mathcal{M} , there is an induced Riemannian metric on $\bar{\mathcal{M}}_\alpha$ so that the projection map π_α in (52) is a so-called Riemannian submersion. Namely, the tangent map $d\pi_\alpha$ at any point $p \in \mathcal{M}_\alpha$ restricts to an isometry between the *horizontal* space \mathcal{H}_p (i.e. vectors perpendicular to the orbit $\bar{p} = K(p)$) and the tangent space of $\bar{\mathcal{M}}_\alpha$ at the point \bar{p} . In this way the total orbit space $\bar{\mathcal{M}}$ in (51) becomes a stratified Riemannian space; in particular, the strata are Riemannian manifolds. Moreover, $\bar{\mathcal{M}}$ has also a global orbital distance metric which measures the distance between orbits in \mathcal{M} (c.f. e.g. [9], [30]). This metric is certainly determined by its restriction to the principal stratum $\bar{\mathcal{M}}_\omega$, which in turn is derived from the Riemannian metric on this stratum. Briefly, the Riemannian manifold $\bar{\mathcal{M}}_\omega$ determines the Riemannian structure on each $\bar{\mathcal{M}}_\alpha$.

Next, we turn to the so-called *slice theorem* which reduces the local study of (K, \mathcal{M}) to linear representation theory, as follows. The *local representation* of the isotropy group $H = K_p$ at a given point p , denoted by $(local)_p$ in (55), is the induced linear action (via differentiation) on the tangent space $T_p\mathcal{M}$ of \mathcal{M} at p . This splits into the *isotropy* and *slice* representation and we write

$$T_p\mathcal{M} = T_pK(p) \oplus \mathcal{H}_p, \quad (local)_p = (Iso)_p + (slice)_p, \quad (55)$$

to indicate that the isotropy representation is the H -action on the tangent space $T_pK(p)$ of the orbit $K(p)$, and the slice representation is the induced action on some H -invariant complementary subspace \mathcal{H}_p (which exists since H is compact). For example, when \mathcal{M} is Riemannian we take \mathcal{H}_p to be the horizontal space, that is, the orthogonal complement of $T_pK(p)$.

Now, the slice theorem says the orbit $K(p)$ has a K -invariant *tubular neighborhood* which has the structure of a *twisted product* $K \times_H \mathcal{H}_p$, consisting of equivalence classes $[k, v]$ in $K \times \mathcal{H}_p$ modulo the relation $[k, v] = [kh^{-1}, hv]$, for $k \in K, h \in H, v \in \mathcal{H}_p$. The "tube" $K \times_H \mathcal{H}_p$ has the structure of a vector bundle over $K(p) \simeq K/H$ with fiber \mathcal{H}_p , and moreover, as a K -space it has the following left action

$$k'[k, v] = [k'k, v], \quad k' \in K$$

and consequently the isotropy groups at points in the tubular neighborhood of $K(p)$ looks like

$$K_{[k, v]} = kH_vk^{-1}$$

In particular, by calculating the slice representation (K_p, \mathcal{H}_p) as the difference

$$(slice)_p = (local)_p - (Iso)_p \quad (56)$$

one can determine the orbit types that occur in the neighborhood. Moreover, there is the following isomorphism of orbit spaces

$$(K \times_{K_p} \mathcal{H}_p)/K \simeq \mathcal{H}_p/K_p \quad (57)$$

which enables one to analyse the local smooth structure of orbit spaces by successive application of the slice theorem.

Finally, we recall the notion of a *fundamental domain* for (K, \mathcal{M}) , or equivalently a *cross-section* for the orbit map (51). This is a closed subset $\Sigma \subset \mathcal{M}$ which intersects each orbit in a unique point, and consequently there is an identification

$$\Sigma \simeq \bar{\mathcal{M}} \quad (58)$$

In the special case that \mathcal{M} is Riemannian and Σ is an orthogonal cross-section, and hence the submanifold Σ is also perpendicular to each orbit, the identification (58) is in fact an isometry. We remark, however, that fundamental domains in the above strict sense exist only in special cases – but fortunately including those orthogonal transformation groups that we shall investigate below.

3.1.2 Compact linear groups on Euclidean spaces

Now, let us consider pairs (K, V) and the corresponding orbit map

$$\pi : V \rightarrow \bar{V} = V/K \quad (59)$$

where K is a compact Lie group acting orthogonally on a Euclidean space V . Since the orbits are compact it is a well known fact that they can be separated by K -invariant polynomial functions \mathfrak{p}_i . Namely, we can choose a Hilbert basis or "sufficiently many" separating and invariant polynomial functions \mathfrak{p}_i as the components of a map

$$\mathfrak{p} = (\mathfrak{p}_1, \mathfrak{p}_2, \dots, \mathfrak{p}_N) : V \rightarrow \mathbb{R}^N$$

which induces an embedding

$$\bar{\mathfrak{p}} : \bar{V} \rightarrow \mathfrak{p}(V) \subset \mathbb{R}^N \quad (60)$$

and hence identifies the orbit space with its image $\mathfrak{p}(V)$. The latter is a semi-algebraic subset, that is, it is defined by polynomial identities and inequalities in \mathbb{R}^N .

Concerning smoothness, we recall that \bar{V} has, on the one hand, the structure of a *differential space* such that a function f on \bar{V} is "smooth" if and only if the composition $f \circ \pi$ is a smooth function on V . On the other hand, there is another approach defining a function on $\mathfrak{p}(V)$ to be "smooth" if it is the restriction of a smooth function on \mathbb{R}^N . Fortunately, the two notions of smoothness are identical, and therefore we may regard the bijective correspondence $\bar{\mathfrak{p}}$ in (60) to

be a diffeomorphism, cf. Schwarz [24]. Consequently, f is "smooth" if and only if $f \circ \pi = \bar{f} \circ \mathfrak{p}$ for some smooth function \bar{f} on \mathbb{R}^N . Note, however, the orbital distance metric on $\bar{V} \simeq \mathfrak{p}(V)$ is certainly not induced from the Euclidean space in (60).

In the Euclidean space V *size* and *shape* are quantities preserved by orthogonal transformations. The norm $\rho = |X|$ is a size function which measures the distance from the origin O , namely the arc-length along rays emanating from O . A nonzero vector can be uniquely scaled to unit size by a homothety transformation $X \rightarrow cX$ ($c > 0$), and by expressing the Euclidean metric ds^2 of V in *polar coordinates* we actually describe V as a Riemannian cone over its unit sphere V^1 :

$$(V^1, d\theta^2) \subset (V, ds^2), \quad ds^2 = d\rho^2 + \rho^2 d\theta^2 \quad (61)$$

where $d\theta^2$ denotes the induced spherical metric on V^1 .

On the other hand, the K -orbit space \bar{V} and its subspace $V^* = V^1/K$ are stratified Riemannian spaces as explained above. Moreover, since the norm function ρ is K -invariant it is also a size function on \bar{V} and then $V^* = (\rho = 1)$ becomes the unit distance "sphere". Furthermore, in analogy with (61) \bar{V} inherits the structure of a stratified Riemannian cone over its "sphere" V^* :

$$(V^*, d\sigma^2) \subset (\bar{V}, d\bar{s}^2), \quad d\bar{s}^2 = d\rho^2 + \rho^2 d\sigma^2 \quad (62)$$

where $d\sigma^2$ denotes the orbital distance metric induced from $d\theta^2$. In this cone the rays emanating from the cone vertex (or base point), still denoted by O , are also geodesics for the metric $d\bar{s}^2$, and homothety transformations move points $\neq O$ along these rays.

3.1.3 The configuration space for m-body systems in d-space

Henceforth, we take V to be the matrix space $M = M(d, m)$ with the orthogonal action of $G = O(d) \times O(m)$ as in (49). By regarding M as the configuration space for an m-body system in \mathbb{R}^d , in analogy with the "physical world" case $d = 3$ described in Section 1, $G_1 = O(d)$ and $G_2 = O(m)$ play the roles of congruence and internal symmetry groups respectively. Then there is the following two-step orbital decomposition of stratified Riemannian spaces

$$M \rightarrow \bar{M} = \frac{M}{G_1} \rightarrow \frac{\bar{M}}{G_2} = \frac{M}{G} \quad (63)$$

where the *internal configuration space* $\bar{M} = \bar{M}(d, m)$ and its symmetry group constitute a natural equivariant system (G_2, \bar{M}) . Again, the metric (62) on \bar{M} will be referred to as the *kinematic metric*, since in the physical case $d \leq 3$ it is, in fact, representing the internal kinetic energy

$$\bar{T} = \frac{1}{2} d\bar{s}^2 = T - T^\omega = \frac{1}{2} ds^2 - T^\omega,$$

namely the total kinetic energy T minus the purely rotational kinetic energy T^ω of the n-body system, see e.g. [30].

The internal symmetry orbit space \bar{M}/G_2 , which is a cone in \mathbb{R}^μ , $\mu = \min\{d, m\}$, is actually a Weyl chamber of Cartan type B_μ . Indeed, the G -representation on M is the isotropy representation of the symmetric space $O(d+m)/O(d) \times O(m)$, which is of type B_μ , and as such it is well understood in terms of the classical invariant theory of matrices and canonical forms modulo orthogonal transformation groups.

Let M^1 be the unit sphere of M . In view of the cone structure (61), (62), we may faithfully replace (63) by its unit distance "sphere" version

$$M^1 \rightarrow M^* = \frac{M^1}{G_1} \rightarrow \frac{M^*}{G_2} = \frac{M^1}{G} = M^{**} \quad (64)$$

where $M^* = M^*(d, m)$ will be referred to as the *shape space*. Its G_2 -orbit space M^{**} is a spherical simplex and it is explicitly described in (85) below.

The orbit types of the above transformation groups will be described in detail. In principle, the topological structure of the shape space M^* can be reconstructed and analyzed as the union of well understood orbit bundles, but in this paper we shall only describe the global topology of M^* in those few cases where it is a manifold. In particular, the cases $d = m - 1$ or $d = m$ are rather simple since (G_2, M^*) is essentially isomorphic to a *linear model*, that is, a sphere S^N or disk D^{N+1} with an orthogonal action of G_2 , see Section 3.4.3. Before turning to these topics, let us have a closer look at some important classical matrix spaces and their natural transformation groups.

3.2 Transformations of the space of symmetric matrices

For $d \leq m$, consider the following sets of $m \times m$ -matrices

$$\begin{array}{ccccccc} \text{Sym}^+(m)_{\leq d} & \subset & \text{Sym}^+(m) & \subset & \text{Sym}(m) & \subset & M(m, m) \\ \cup & & \cup & & \cup & & \\ B^+(d) \simeq B^+(m)_{\leq d} & \subset & B^+(m) & \subset & B(m) & \subset & \text{Diag}(m) \simeq \mathbb{R}^m \end{array} \quad (65)$$

where in the first row $\text{Sym}(m)$ is the space of symmetric matrices, $\text{Sym}^+(m)$ is the convex cone of positive semidefinite matrices, and $\text{Sym}^+(m)_{\leq d}$ is the subcone of matrices of rank at most d . In the second row $\text{Diag}(m)$ is the space of diagonal matrices $\text{diag}(\lambda_1, \dots, \lambda_m)$, containing the cone

$$B(m) = \{Y = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_m); \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m\}, \quad (66)$$

and the subcones $B^+(m)$ and $B^+(m)_{\leq d}$ are given by the additional constraint $\lambda_m \geq 0$, and $\lambda_i = 0$ for all $i > d$, respectively.

With the usual inner product on $M(m, m)$

$$Y_1 \cdot Y_2 = \text{trace}(Y_1^t Y_2), \quad \text{cf. (14)}$$

where Y^t means the transpose of Y , $O(m)$ acts orthogonally by conjugation

$$(\varphi, Y) \rightarrow \varphi Y \varphi^{-1}, \quad \varphi \in O(m), Y \in M(m, m) \quad (67)$$

and the subsets in the first row of (65) are clearly invariant. Thus the transformation group $(O(m), \text{Sym}(m))$ is the second symmetric tensor product $S^2 \rho_m$

of the standard representation $\rho_m = (O(m), \mathbb{R}^m)$, and let us recall the classical result about diagonalization of symmetric matrices. Namely, every symmetric matrix is conjugate to a unique matrix - its *canonical form* - in the cone $B(m)$, which is therefore a fundamental domain (or cross section) of the orbit map

$$Sym(m) \rightarrow \frac{Sym(m)}{O(m)} \simeq \frac{\mathbb{R}^m}{S_m} \simeq B(m) \subset Sym(m) \quad (68)$$

As indicated, $B(m)$ is also a fundamental domain (or cross section) for (S_m, \mathbb{R}^m) , where S_m is the symmetric group acting on \mathbb{R}^m by permuting the coordinates λ_i .

Fundamental domains for the $O(m)$ -action restricted to invariant subspaces of $Sym(m)$ are the corresponding subsets of $B(m)$ in the second row of (65), consequently

$$B^+(m) \simeq \frac{\mathbb{R}^m}{B_m} \simeq \frac{Sym^+(m)}{O(m)} \supset \frac{Sym^+(m)_{\leq d}}{O(m)} \simeq \frac{\mathbb{R}^d}{B_d} \simeq B^+(d) \quad (69)$$

As indicated, we recognize $B^+(k)$ as a Weyl chamber for the compact connected Lie group $SO(2k+1)$, namely a fundamental domain of the canonical representation of the Weyl group $B_k \supset S_k$ acting on \mathbb{R}^k as a group generated by reflections.

Note that $Sym(m)$ is not an irreducible $O(m)$ -space. In fact, there is the orthogonal and invariant decomposition

$$Sym(m) = Sym^0(m) \oplus \mathbb{R}^1 \quad (70)$$

where the trivial summand \mathbb{R}^1 is spanned by the identity matrix, and the first summand, which consists of matrices of trace zero, gives the irreducible representation $S^2\rho_m - 1$ with fundamental domain $B^0(m) \subset B(m)$ defined by $\sum \lambda_i = 0$:

$$B^0(m) \simeq \frac{\mathbb{R}^{m-1}}{S_m} \simeq \frac{Sym^0(m)}{O(m)} \quad (71)$$

Here we have also recognized $B^0(m)$ as a fundamental domain of the canonical representation of the Weyl group $A_{m-1} = S_m$ of $SU(m)$.

In addition to its Euclidean norm where $|Y| = \sqrt{\text{trace}(Y^2)}$, the positive cone $Sym^+(m)$ has another $O(m)$ -invariant size function $\rho^*(Y) = \text{trace}(Y)$, with corresponding "unit sphere"

$$Sym^*(m) = \{Y \in Sym^+(m) ; \text{trace}(Y) = 1\} \quad (72)$$

For later reference, we note that its intersection with the Euclidean unit sphere in $Sym(m)$ is the subset

$$Sym^*(m) \cap Sym(m)^1 = Sym^*(m)_1 = \{Y \in Sym^*(m) ; rk(Y) = 1\} \quad (73)$$

Let $Y \in B(m)$. The tangent space at Y of the $O(m)$ -orbit through Y in $Sym(m)$ is the subspace

$$[so(m), Y] = \{SY - YS ; S \text{ is skew-symmetric}\} \subset Sym(m),$$

and it is perpendicular to $B(m)$ since the inner product with any $Y' \in B(m)$ is

$$(SY - YS) \cdot Y' = \text{trace}((SY - YS)Y') = 0$$

A direct consequence of this observation is as follows :

Remark 3.1 *The fundamental domain $B(m)$ of $(O(m), \text{Sym}(m))$ is an orthogonal cross section, in the sense that it is perpendicular to the $O(m)$ -orbits. Therefore, the orbit space of $\text{Sym}(m)$ (with the orbital distance metric) is isometric to $B(m)$. Consequently, orbit spaces of various $O(m)$ -invariant subspaces of $\text{Sym}(m)$ in (65) are isometric to the corresponding subsets of $B(m)$.*

Finally, let us determine the orbit types of the $O(m)$ -action on symmetric matrices. To this end, consider a matrix $Y \in B(m)$ and let the "strings" of equal entries λ_i in (66) have length m_1, m_2, \dots , so that

$$m = m_1 + m_2 + \dots + m_p \quad (m_i > 0, p > 0)$$

is a partition of m and p is the number of different λ_i 's. It is easy to see that the isotropy group at Y is the "block" orthogonal matrix group

$$O(m_1, m_2, \dots, m_p) = O(m_1) \times O(m_2) \times \dots \times O(m_p) \subset O(m) \quad (74)$$

and therefore the partition function $\pi(m)$ of m enumerates the different orbit types. The orbits are in fact connected and therefore they are also $SO(m)$ -orbits, that is,

$$\frac{\text{Sym}(m)}{SO(m)} = \frac{\text{Sym}(m)}{O(m)} \simeq B(m)$$

3.3 Algebraic realization of orbit spaces and orbital stratification

With the results from the previous subsection as a basis, the construction of the orbit spaces in (63) or (64) as semialgebraic subsets is based upon the following two simple but fundamental properties of matrices :

- The polynomial map

$$\mathbf{p} : M(d, m) \rightarrow \text{Sym}(m), \quad X \rightarrow Y = X^t X \quad (75)$$

preserves the matrix rank, and its image consists of all positive semidefinite symmetric matrices Y of rank $\leq \min\{d, m\}$. Moreover, for $\psi \in O(d), \varphi \in O(m)$,

$$\mathbf{p}(\psi X \varphi^{-1}) = \varphi \mathbf{p}(X) \varphi^{-1} = \varphi Y \varphi^{-1} \quad (76)$$

and the entries of Y are the inner products $\mathbf{x}_i \cdot \mathbf{x}_j$ of the column vectors \mathbf{x}_i of X .

- Consider a finite collection of vectors \mathbf{x}_j in a Euclidean space \mathbb{R}^d . Then the collection is uniquely determined, up to $O(d)$ -congruence, by the inner products $\mathbf{x}_i \cdot \mathbf{x}_j$. For a proof we refer to Weyl[31], page 52.

The map \mathfrak{p} in (75) is clearly constant on each $O(d)$ -orbit in $M = M(d, m)$ and induces a surjective map

$$\bar{\mathfrak{p}} : \bar{M}(d, m) \rightarrow \text{Sym}^+(m)_{\leq d} \subset \text{Sym}(m) \quad (77)$$

which is, moreover, injective by the above mentioned result of Weyl. In other words, the components of \mathfrak{p} , namely the collection of inner products $\mathbf{x}_i \cdot \mathbf{x}_j$, constitute a Hilbert basis in the sense of Section 3.1.1. Consequently, $\bar{\mathfrak{p}}$ is a diffeomorphism between differential spaces which identifies $\bar{M} = M/O(m)$ with the set of positive semidefinite symmetric $m \times m$ -matrices of rank at most d . In particular,

$$\bar{M}(d, m) = \bar{M}(m, m) \simeq \text{Sym}^+(m) \text{ if } d \geq m, \quad (78)$$

so we will henceforth assume without loss of generality that $d \leq m$. Thus we have explained why all orbit spaces involved are realizable as appropriate spaces of symmetric matrices, amenable to the setting in Section 3.2.

3.3.1 Cross sections and canonical forms

In general, the inclusions $M(k, m) \subset M(k+1, m)$, defined by taking the last row to be zero, induce the increasing filtration

$$\bar{M}(1, m) \subset \bar{M}(2, m) \subset \dots \subset \bar{M}(m, m) \simeq \text{Sym}^+(m) \quad (79)$$

and this coincides with the matrix rank filtration of the positive cone $\text{Sym}^+(m)$, namely for $k+i \leq m$

$$\bar{M}(k, m) = \bar{M}(k+i, m)_{\leq k} \simeq \text{Sym}^+(m)_{\leq k}$$

On the other hand, $G_2 = O(m)$ acts on the spaces in (77). The action on $\bar{M} = \bar{M}(d, m)$ is induced from the action (49), and G_2 acts on $\text{Sym}(m)$ by conjugation (67). From the equivariance property (76) it follows that these actions commute with the map $\bar{\mathfrak{p}}$, that is, $\bar{\mathfrak{p}}$ is an isomorphism of G_2 -spaces and hence induces a diffeomorphism between G_2 -orbit spaces

$$\bar{\mathfrak{p}} : \frac{M}{G} = \frac{\bar{M}}{G_2} \rightarrow \frac{\text{Sym}^+(m)_{\leq d}}{G_2} \simeq B^+(d) \subset \text{Sym}^+(m)_{\leq d} \simeq \bar{M} \quad (80)$$

The cross section $B^+(d) \subset \bar{M}$ in (80), being transversal to the G_2 -orbits in \bar{M} , further lifts to a cross section $M^+(d)$ of the composite orbit map $M \rightarrow M/G$, as follows :

$$B^+(d) \simeq M^+(d) \subset M : \left\{ \begin{array}{l} X = \text{diag}(r_1, r_2, \dots, r_d) \\ r_1 \geq r_2 \geq \dots \geq r_d \geq 0 \end{array} \right. \quad (81)$$

where $X = (x_{ij})$ has $x_{ii} = r_i$ and zero entries otherwise. Thus we have also established the diagonalization procedure saying that every matrix in $M = M(d, m)$ can be transformed by the action (49) to a unique matrix - its canonical form - in the subset $M^+(d)$. As a (simplicial) Euclidean cone in the $\{r_i\}$ -coordinate space \mathbb{R}^d , $M^+(d)$ is also the Weyl chamber of the Weyl group B_d , cf. (69).

3.3.2 Orthogonality of cross sections

In view of Remark 3.1, the cone $M^+(d)$ in (81) is orthogonal to the G -orbits and is therefore an orthogonal cross section for (G, M) . Hence, we have an isometry $M/G \simeq M^+(d)$ of Euclidean cones with the metric

$$ds^2 = \sum_{i=1}^d dr_i^2 = \frac{1}{4} \sum_{i=1}^d \frac{1}{\lambda_i} d\lambda_i^2 \quad (82)$$

Here $\{r_i\}$ from (81) and $\{\lambda_i\}$ from (66) are coordinate systems for the cone and are related by the polynomial map \mathbf{p} in (75), which restricts to a diffeomorphism

$$\mathbf{p} : M^+(d) \rightarrow B^+(d) \subset \text{Sym}^+(m), \quad r_i \rightarrow r_i^2 = \lambda_i \quad (83)$$

Thus, the G -orbit of a given matrix $X \in M = M(d, m)$ has coordinates λ_i interpreted as the eigenvalues of $X^t X$. In the special case that $d = m$ and X is symmetric, the numbers r_i are the absolute values of the eigenvalues of X .

Remark 3.2 *At this point, observe that the internal configuration space $\bar{M} = \bar{M}(d, m)$ has two metrics, namely the kinematic metric (i.e. the $O(d)$ -orbital distance metric) and the induced "fake" metric as a subset $\text{Sym}^+(m)_{\leq d}$ of the Euclidean space $\text{Sym}(m)$, cf. (77). For both metrics the symmetry group $O(m)$ is actually an isometric transformation group with $B^+(d)$ as a fundamental domain (i.e. cross section). In the "fake" metric $B^+(d)$ inherits the Euclidean metric $\sum d\lambda_i^2$ and is by Remark 3.1 an orthogonal cross section, whereas in the kinematic (but still Euclidean) metric (82) of \bar{M} the orthogonality property of $B^+(d)$ fails. See also Remark 3.4 below.*

3.3.3 Rank and subrank stratification of matrix spaces

The *subrank stratification* of the matrix space $M = M(d, m)$ is a natural refinement of the usual *rank stratification*. The latter coincides with the orbit type stratification of the action of G_1 (or G_2) on M and the mentioned refinement coincides with the orbit type stratification of the action of the full group

$$G = G_1 \times G_2 = O(d) \times O(m)$$

In order to describe the combinatorial structure involved we turn to the two-step orbital decomposition (64), where the first orbit space is the shape space

$$M^* = M^*(d, m) \simeq \text{Sym}^*(m)_{\leq d}, \quad \text{cf. (72)} \quad (84)$$

and the final G -orbit space is

$$M^{**} = M^{**}(d, m) = M^*/G_2 = M^1/G$$

which we may identify with the following fundamental domain in the sphere M^1 :

$$\Delta^{d-1} = M^+(d) \cap M^1 : r_1 \geq r_2 \geq \dots \geq r_d \geq 0, \sum r_i^2 = 1 \quad (85)$$

This is a spherical simplex whose structure, indeed, identifies it with the spherical Weyl chamber of type B_d (see (69)); in particular, it is homeomorphic to the closed disk D^{d-1} .

Thus we also conclude that the above simplex (85) is a cross section for the transformation group (G, M^1) , and hence any matrix in M^1 is G -equivalent to a unique matrix - its *canonical form* - in the simplex. On the other hand, a typical matrix of rank k in (85) may be characterized by the following numerical data

$$X : \begin{cases} r_1 = r_{k_1} > r_{k_1+1} = r_{k_1+k_2} > r_{k_1+k_2+1} = \dots = r_k > 0 = r_{k+1} \\ 1 \leq \sum_{i=1}^p k_i = k \leq d, \quad k_i > 0, \quad p > 0, \end{cases} \quad (86)$$

and then its *subrank* is defined to be the corresponding (unordered) partition of k into p positive integers

$$\kappa = (k_1, k_2, \dots, k_p), \quad |\kappa| = k_1 + \dots + k_p = k \quad (87)$$

which records the length of the strings of equal numbers in (86).

We label the rank (resp. subrank) stratum by the subscript k (resp. κ) of the corresponding sets. Thus X belongs to M_k^1 if and only if its G_1 -orbit has the type $O(d-k)$ (cf. also (100) below). Moreover, by (85) the image of M_k^1 in M^{**} is the "semi-open" simplicial disk

$$M_k^{**} \simeq \Delta_k^{k-1} \subset \Delta_k^{d-1} : r_1 \geq r_2 \geq \dots \geq r_k > r_{k+1} = 0, \sum r_i^2 = 1,$$

which is subdivided into its various subrank strata

$$M_\kappa^{**} \simeq \Delta_\kappa^{k-1} \subset \Delta_\kappa^{d-1} \subset \Delta^{d-1} \simeq M^{**} \quad (88)$$

In particular, the principal stratum has subrank $\kappa_0 = (1, 1, \dots, 1)$ with $|\kappa_0| = d$, and $M_{\kappa_0}^{**} \simeq \Delta_{\kappa_0}^{d-1}$ is the interior of the spherical simplex Δ^{d-1} and is therefore homeomorphic to \mathbb{R}^{d-1} .

To see why the subrank stratification actually coincides with the G -orbit type stratification, consider the isotropy group G_X of the matrix in (86), using the following notation for groups :

$$\begin{aligned} \Delta H &= \text{diag}(H \times H) \simeq H \quad (\text{diagonal embedding}) \\ O(k_1, \dots, k_p) &= O(k_1) \times \dots \times O(k_p) \subset O(k) \\ G(k_1, \dots, k_p) &= O(d-k) \times \Delta O(k_1, \dots, k_p) \times O(m-k) \subset G \end{aligned} \quad (89)$$

Then it is not difficult to show by direct calculations with matrices that

$$G_X = \{(g_1, g_2) \in O(d) \times O(m) ; g_1 X g_2^{-1} = X\} = G(k_1, \dots, k_p) \quad (90)$$

(cf. Table 1, #10, in [27]), and hence the isotropy types uniquely characterize the subrank strata, as claimed.

We are particularly interested in the stratification of M^* , but now there are two natural options, namely the induced subrank stratification and the G_2 -orbit type stratification. Our next claim, however, is that they are identical,

and hence one can describe M^* as a union of G_2 -orbit bundles lying over the various strata (88).

To calculate the G_2 -orbit types and how they correspond to the subrank strata we proceed as follows. Let $X^* \in M^*$ be the image of a matrix X of rank k and consider the "large" orbit $G(X) \subset M$. As a G_1 -space all the orbits of $G(X)$ have the same type as the "small" orbit $G_1(X)$. Moreover, the orbit space $G(X)/G_1 \subset M^*$ coincides with the G_2 -orbit through X^* , namely the set $G_2(X^*)$. Thus, the "large" G -orbit through X has the structure of a G_1 -orbit bundle, and in terms of homogeneous spaces (i.e. coset spaces of groups) this fiber bundle can be described as follows :

$$\begin{array}{ccccc} G_1(X) & \rightarrow & G(X) & \rightarrow & G_2(X^*) \\ \parallel & & \parallel & & \parallel \\ \frac{O(d)}{O(d-k)} & \rightarrow & \frac{G}{G(k_1, k_2, \dots, k_p)} & \rightarrow & \frac{O(m)}{O(k_1, k_2, \dots, k_p) \times O(m-k)} \end{array} \quad (91)$$

The base space $G_2(X^*)$ of the fiber bundle is in turn a fiber of another bundle, namely the multistratum M_κ^* which as a G_2 -orbit bundle fibers as follows

$$\frac{O(m)}{O(k_1, k_2, \dots, k_p) \times O(m-k)} \rightarrow M_\kappa^* \rightarrow M_\kappa^{**} \simeq \Delta_\kappa^{k-1} \quad (92)$$

In this way the subrank stratification of the shape space

$$M^*(d, m) = \bigcup_{k=1}^d M_k^* = \bigcup_{k=1}^d \bigcup_{|\kappa|=k} M_\kappa^*$$

has strata which are fiber bundles of type (92), and there are altogether

$$\pi(1) + \pi(2) + \dots + \pi(d) \quad (93)$$

such bundles (or strata), where $\pi(k)$ is the partition function of k . The principal stratum $M_{\kappa_0}^*$ constitutes an open and dense subset of M^* , and the principal orbit bundle

$$\frac{O(m)}{O(1)^d \times O(m-d)} \rightarrow M_{\kappa_0}^* \rightarrow M_{\kappa_0}^{**} \approx \mathbb{R}^{d-1} \quad (94)$$

is actually trivial (e.g. since the base space is contractible). The orbit type, either for the G_2 -action on M^* or the G -action on M^1 , for a stratum of subrank $\kappa = (k_1, k_2, \dots, k_p)$ can be read off from (91).

Example 3.3 $d = 3$: $M(3, n-1) \simeq M_n$ is the centered configuration space for n -body systems in 3-space. See Section 3.4.4 for a brief description of the shape spaces M_n^* when n is small. In general, for $n > 3$ their G_2 -orbit space

$$\frac{M_n^*}{G_2} = M^{**}(3, n-1) \simeq \Delta^2$$

is the spherical triangle on the sphere $r_1^2 + r_2^2 + r_3^2 = 1$ with vertices

$$A = (1, 0, 0), \quad B = (1/\sqrt{2}, 1/\sqrt{2}, 0), \quad C = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$$

On the circle $r_3 = 0$ lies the closed circular arc of length $\pi/4$

$$[A, B] : r_1 \geq r_2 \geq 0$$

which consists of those points of rank ≤ 2 . It has three strata, namely the vertex A (of rank 1), the vertex B and the open arc (A, B) (both of rank 2). However, the open arcs

$$(A, C) : r_1 > r_2 = r_3 > 0, \quad (B, C) : r_1 = r_2 > r_3 > 0$$

constitute the same stratum of subrank $\kappa = (2, 1)$. Thus, although the triangle decomposes into $7 = 3 + 3 + 1$ strata components, namely vertices, open edges and the interior of Δ^2 , there are only 6 multistrata (since one of them has 2 components), in accordance with the enumeration formula (93).

Remark 3.4 In general, the kinematic metric on the shape space M^* is uniquely determined by its restriction to the principal stratum, which is a Riemannian manifold $(M_{\kappa_0}^*, d\sigma^2)$. There is a general procedure, in the setting of equivariant differential geometry, for the calculation (or description) of the Riemannian connection on such a principal orbit bundle, see for example [4]. The simplest case $d = 3$, as in the above example, is analyzed in Hsiang[9], but similar calculations can also be done for $d > 3$. We shall, however, leave the geometric issues here and concentrate on the topological structures in the sequel.

3.4 Topology of the shape space and related spaces

For fixed m the largest shape space (with $d = m$) is

$$M^*(m, m) \simeq \text{Sym}^*(m) \tag{95}$$

and by (84) it contains any other shape space

$$M^*(d, m) = M^*(m, m)_{\leq d} \simeq \text{Sym}^*(m)_{\leq d}$$

as the union of those strata of matrix rank at most d . Therefore, we shall refer to the space (95) as the m -universal shape space. Thus there is the increasing and $O(m)$ -invariant rank filtration

$$M^*(1, m) \subset M^*(2, m) \subset \dots \subset M^*(m, m) \tag{96}$$

and corresponding $O(m)$ -orbit spaces $\approx \Delta^{d-1}$, $1 \leq d \leq m$,

$$M^{**}(1, m) \subset M^{**}(2, m) \subset \dots \subset M^{**}(m, m) = \Delta^{m-1} \tag{97}$$

where the first three spaces in the chain (97) are, indeed, the inclusions of simplices

$$\{A\} = \Delta^0 \subset [A, B] = \Delta^1 \subset [A, B, C] = \Delta^2$$

from Example 3.3.

By inspection the chain in (96) starts with

$$\begin{aligned} M^*(1, m) &= S^{m-1}/O(1) = \mathbb{R}P^{m-1} \\ M^*(2, m) &= S^{2m-1}/O(2) = \mathbb{C}P^{m-1}/\mathbb{Z}_2 \end{aligned} \quad (98)$$

where

$$\mathbb{R}P^{m-1} \subset \mathbb{C}P^{m-1} = S^{2m-1}/SO(2)$$

are the real and complex the projective $(m-1)$ -space, and $\mathbb{Z}_2 = O(2)/SO(2)$ acts on $\mathbb{C}P^{m-1}$ by complex conjugation with $\mathbb{R}P^{m-1}$ as fixed point set. In particular, the vertex $\Delta^0 = M^{**}(1, m)$ of the simplex Δ^{m-1} in (97) is the single $O(m)$ -orbit $M^*(1, m)$, with the topology of a real projective space as noted above.

The topology of the next shape space $M^*(2, m)$ is more difficult to describe. One approach is to utilize the fact that it is a $O(m)$ -space of *cohomogeneity one*, that is, with one-dimensional orbit space. This holds since

$$\frac{M^*(2, m)}{O(m)} = M^{**}(2, m) \simeq \Delta^1 = [A, B]$$

is an interval. We also know the orbit types $(K_1), (K_2), (H)$ corresponding to the three strata A, B and (A, B) respectively. Therefore, the space can be described topologically in terms of its orbit types by the construction (cf. e.g. [29], Vol.1, Chap. IV) known as the *equivariant union*

$$M^*(2, m) = M(\pi_1) \cup M(\pi_2)$$

of the mapping cylinders $M(\pi_i)$ of the canonical projections $\pi_i : O(m)/H \rightarrow O(m)/K_i$, where in our case

$$\begin{aligned} K_1 &= O(1) \times O(m-1), \quad K_2 = O(2) \times O(m-2), \\ H &= K_1 \cap K_2 = O(1)^2 \times O(m-2) \end{aligned} \quad (99)$$

The space $M^*(2, 2)$ is a 2-disk and $M^*(2, 3)$ is actually homeomorphic to S^4 , see (104) and Section 3.4.4). For $m > 3$ $M^*(2, m)$ fails to be a manifold in a neighborhood of the singular $O(m)$ -orbit

$$A = O(m)/K_1 = M^*(1, m) \simeq \mathbb{R}P^{m-1},$$

but two copies of $M^*(2, m)$ glued together along A yields, indeed, the differentiable manifold $S^{2m-1}/SO(2) = \mathbb{C}P^{m-1}$. Below we will return to this construction and further investigate the topology of the spaces $M^*(d, m)$ for $2 < d \leq m$.

3.4.1 Local and global topology

By referring to Section 3.1 concerning compact transformation groups and the slice theorem, let us first recall a nice property of *regular* representations of classical groups such as $O(d)$. Namely, we consider the linear action of $O(d)$ on \mathbb{R}^N by some representation of type $\Phi = m\rho_d + \tau_q$, $N = md + q$, which is m

copies of the standard representation ρ_d plus a q -dimensional trivial summand τ_q . Then the orbit types constitute the following "connected" string of *regular* subgroups

$$O(k) ; \quad d - m \leq k \leq d \quad (100)$$

where $O(k) = 1$ if $k \leq 0$. To calculate the slice representation Φ_k at a point with isotropy group $O(k)$ we calculate the difference as in (56), with the local representation equal to Φ restricted to $O(k)$, that is, $m\rho_k + (\text{trivial})$. Moreover, the isotropy representation $Iso(d, k)$ of $O(k)$ is the "linearized" action on the Stiefel manifold $O(d)/O(k) = SO(d)/SO(k)$ at the base point, and a simple calculation (involving the adjoint representation of $O(d)$) yields the representation $(d - k)\rho_k + (\text{trivial})$, consequently

$$\Phi_k = \Phi|_{O(k)} - Iso(d, k) \equiv (m - d + k)\rho_k \pmod{\text{trivial}} \quad (101)$$

From this we conclude that each slice representation inherits the *regularity* property of the original action Φ .

Now we turn to the "left side" action of $O(d)$ on the matrix space $M = M(d, m) \simeq \mathbb{R}^{md}$, which is just the regular representation $\Phi = m\rho_d$. However, we shall rather consider the restricted action on the unit sphere M^1 , noting that the only change in the above calculations is that the local representations lose a trivial summand τ_1 . The orbit types still constitute the string (100) and also (101) holds, except that $k < d$ in (100) since $O(d)$ has no fixed point on the sphere M^1 .

Lemma 3.5 (i) *The m -universal shape space $M^*(m, m) \simeq \text{Sym}^*(m)$ is a manifold with boundary*

$$\partial M^*(m, m) = M^*(m - 1, m) \simeq \text{Sym}^*(m)_{\leq m-1}$$

(ii) *For $p > 0$, $M^*(d, d + p)$ is a manifold (and with no boundary) if and only if $p = 1$ or $d = 1$.*

Proof. By (57), the local topology of $M^*(m, m)$ around an $O(m)$ -orbit of type $O(k)$ is the topology of the orbit space of the slice representation $(O(k), \Phi_k)$ in M^1 , with $m = d$ in (101). Hence, part (i) of the lemma follows by induction on $m \geq 1$, since at the initial step $m = 1$ the orbit space of $(O(1), \rho_1)$ is the half-line $\mathbb{R}_+^1 = [0, \infty)$.

For the proof of (ii) we may assume $d > 1$, and the slice representation (101) of $O(k)$ is $\Phi_k = (k + p)\rho_k \pmod{\text{trivial}}$. By induction the proof reduces to the crucial singular case $(O(1), (1 + p)\rho_1)$, where $O(1)$ acts by inversion $v \rightarrow -v$, and here the orbit space is the cone $\frac{\mathbb{R}^{p+1}}{O(1)} = C(\mathbb{R}P^p)$ over the real projective p -dimensional space. This is a manifold if and only if $p = 1$, in which case the orbit space is homeomorphic to the Euclidean plane

$$\frac{\mathbb{R}^2}{O(1)} = C(\mathbb{R}P^1) \approx C(S^1) = \mathbb{R}^2 \quad (102)$$

■

Next, let us actually determine the topological type of the compact manifold $M^*(m, m) \simeq \text{Sym}^*(m)$ and its boundary. By the $O(m)$ -action on $M(m, m)$ every $X = (x_{ij})$ can be mapped to the subset

$$T^+(m) \subset M(m, m)$$

of upper triangular matrices with diagonal entries $x_{ii} \geq 0$, and therefore the restriction of the polynomial map in (75)

$$\mathbf{p} : T^+(m) \rightarrow \text{Sym}^+(m) \quad (103)$$

is still surjective. The subset $T^+(m)_m$ of matrices X with all $x_{ii} > 0$, i.e. of maximal rank m , is clearly diffeomorphic to the Euclidean space $\mathbb{R}^{m(m+1)/2}$, and it is easy to verify that different matrices X lie on different $O(m)$ -orbits. Therefore the following lemma must hold.

Lemma 3.6 *The polynomial map in (103) further restricts to a diffeomorphism*

$$\mathbf{p} : T^+(m)_m \simeq \text{Sym}^+(m)_m \simeq \mathbb{R}^{m(m+1)/2}$$

with the set of positive definite symmetric matrices. Hence, the interior of $\text{Sym}^(m)$ is the open "sphere-octant"*

$$\text{Sym}^*(m)_m \simeq T^+(m)_m \cap M^1(m, m) : \sum x_{ij}^2 = 1, x_{ii} > 0$$

which is diffeomorphic to $\mathbb{R}^{m(m+1)/2-1}$.

We know from the previous two lemmas that $\text{Sym}^*(m)$ is a compact manifold with boundary and its interior is an open disk. Using some manifold theory this information actually suffices to conclude that $\text{Sym}^*(m)$ is a closed disk. For example, one may apply the so-called collar neighborhood theorem and the h-cobordism theorem, cf. Milnor[22]). Anyhow, we have the topological types

$$\begin{aligned} M^*(m, m) &\simeq \text{Sym}^*(m) \approx D^{m(m+1)/2-1} \\ M^*(m-1, m) &\simeq \text{Sym}^*(m)_{\leq m-1} \approx S^{m(m+1)/2-2} \end{aligned} \quad (104)$$

and taking the cone over these spaces yields the topological types

$$\begin{aligned} \bar{M}(m, m) &\approx \mathbb{R}^{m(m+1)/2-1} \times [0, \infty) \quad (\text{half-space}) \\ \bar{M}(m-1, m) &\simeq \partial \bar{M}(m, m) \approx \mathbb{R}^{m(m+1)/2-1} \quad (\text{Euclidean space}) \end{aligned}$$

Remark 3.7 *The linear model construction in Section 3.4.3 provides another proof of the homeomorphisms (104).*

3.4.2 Branched coverings and generalized Hopf fibrations

The usage of $SO(d)$ rather than $O(d)$ as the congruence group acting on $M(d, m)$ may lead to a different "shape space", and we shall explain this distinction below. First, observe that for $k \geq 1$ we have connected and hence equal orbits

$$O(d)/O(k) \simeq SO(d)/SO(k)$$

and consequently the orbit spaces of $SO(d)$ and $O(d)$ will be identical as long as $d > m$, see (100). However, due to (78) we have been assuming $d \leq m$ and therefore $k = 0$ also occurs in the string (100), so the two orbit spaces cannot be identical. In order to explain the difference the following discussion will be helpful.

Consider a space \tilde{Q} with a given involution σ (i.e. a transformation of order two) and hence we shall regard $\mathbb{Z}_2 = \{Id, \sigma\}$ as a transformation group on \tilde{Q} . We denote the orbit space of \mathbb{Z}_2 by Q and the fixed point set by Σ . We also assume $\Sigma \neq \emptyset$ and regard it as a subset of both \tilde{Q} and Q . Then the orbit map

$$\pi : \tilde{Q} \rightarrow \frac{\tilde{Q}}{\mathbb{Z}_2} = Q \quad (105)$$

is an example of a *ramified double covering* which is ramified along Σ . Conversely, starting from Q and Σ we may reconstruct \tilde{Q} as the *double* of Q

$$2Q = Q \cup_{\Sigma} Q \simeq \tilde{Q}$$

by taking two copies of Q and identify (or glue together) their "singular" set Σ . However, the pair (Q, Σ) does not always lead to a unique double space, so we shall rather have in mind a specified ramified covering as in (105). Thus, we define for $d \leq m$

$$2M^*(d, m) = M^*(d, m) \cup_{M^*(d-1, m)} M^*(d, m) \quad (106)$$

and refer to the orbit map

$$2M^*(d, m) \simeq \frac{M^1(d, m)}{SO(d)} \rightarrow \frac{M^1(d, m)}{O(d)} = M^*(d, m) \quad (107)$$

of the action of $O(d)/SO(d) = \mathbb{Z}_2$, which is, indeed, a double covering ramified along the fixed point set $M^*(d-1, m)$ of \mathbb{Z}_2 .

Lemma 3.8 *For $p \geq 0$ and $d > 1$, the space $2M^*(d, d+p)$ is a manifold (and with no boundary) if and only if $p = 0$ or $d = 2$.*

Proof. For $d = 2$ the $SO(2)$ -orbit space is the complex projective space $\mathbb{C}P^{p+1}$, so let us assume $d > 2$. Using again the slice theorem and an inductive argument, the crucial case will be the topology of the orbit space of $(SO(2), (p+2)\rho_2)$, namely the cone

$$\frac{\mathbb{R}^{2(2+p)}}{SO(2)} = C\left(\frac{S^{2(2+p)-1}}{SO(2)}\right) = C(\mathbb{C}P^{p+1})$$

over the complex projective space $\mathbb{C}P^{p+1}$. It is well known that this is a manifold (i.e. locally Euclidean) if and only if $p = 0$, in which case it is a cone homeomorphic to Euclidean 3-space

$$C(\mathbb{C}P^1) \approx C(S^2) = \mathbb{R}^3 \quad (108)$$

■

Lemma 3.9 *The induced smooth functional structure on the cone (102) (resp. (108)) is a refinement of the standard differentiable structure of \mathbb{R}^2 (resp. \mathbb{R}^3). The structures are identical away from the origin (cone vertex), where the cone fails to be a smooth manifold.*

Proof. Consider the case (108), namely the orbit space of $(SO(2), 2\rho_2)$ acting on $\mathbb{R}^2 \times \mathbb{R}^2$. To describe it algebraically, let \mathbf{u}, \mathbf{v} denote vectors in \mathbb{R}^2 and define the polynomial map

$$\mathbf{p} = (X, Y, Z, W) : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}^4, \text{ see (60)}$$

where $X = |\mathbf{u}|^2 - |\mathbf{v}|^2, Y = 2\mathbf{u} \cdot \mathbf{v}, Z = 2(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{k}, W = |\mathbf{u}|^2 + |\mathbf{v}|^2$, which identifies the orbit space with the cone in 4-space

$$X^2 + Y^2 + Z^2 = W^2, \quad W \geq 0$$

By projecting the cone onto the coordinate (X, Y, Z) -space we identify it with \mathbb{R}^3 as in (108), with (standard) coordinate functions X, Y, Z which are certainly smooth in the induced orbital functional structure. However, at the origin the function $f = \sqrt{X^2 + Y^2 + Z^2}$ is not smooth with respect to the standard structure, although its composition with the orbit map is by our construction the smooth function W . The case (102) is similar but simpler, using $\mathbf{p} = (X, Y, W)$ and $(u, v) \in \mathbb{R}^1 \times \mathbb{R}^1$. ■

According to (104) and (107) we have constructions of orbital fibrations of spheres over spheres involving two types of orthogonal transformation groups, as follows :

Case 1 : Consider the regular representation $(O(m-1), m\rho_{m-1})$, acting on the unit sphere $M^1(m-1, m) = S^{(m-1)m-1}$ with orbit space as in the second line of (104),

$$S^{(m-1)m-1} \rightarrow \frac{S^{(m-1)m-1}}{O(m-1)} \approx S^{m(m+1)/2-2}, \quad m = 2, 3, 4, \dots \quad (109)$$

Here the initial case $m = 2$ is the trivial Hopf fibration or double covering

$$O(1) \hookrightarrow S^1 \xrightarrow{\pi} \mathbb{R}P^1 = \frac{S^1}{O(1)} \simeq S^1 \quad (110)$$

Case 2 : Consider the regular representation $(SO(m), m\rho_m)$, acting on the unit sphere $M^1(m, m) = S^{m^2-1}$. The orbit space coincides with the "double" construction (106) applied to the first line of (104)

$$\frac{M^1(m, m)}{SO(m)} = 2M^*(m, m) \approx 2D^{m(m+1)/2-1} = S^{m(m+1)/2-1}$$

and there is the orbit map

$$S^{m^2-1} \rightarrow \frac{S^{m^2-1}}{SO(m)} \approx S^{m(m+1)/2-1}, \quad m = 2, 3, 4, \dots \quad (111)$$

which for $m = 2$ is the well known Hopf fibration

$$SO(2) \hookrightarrow S^3 \xrightarrow{\pi} \mathbb{C}P^1 = \frac{S^3}{SO(2)} \simeq S^2 \quad (112)$$

Remark 3.10 *With the induced differential structure the above quotient spheres are, of course, stratified differentiable manifolds (cf. Section 3.1) which are also locally Euclidean spaces since they are spheres in the topological sense. In fact, for $m = 2$ they are the standard sphere (i.e. S^1 or S^2) in the differentiable sense as well. However, for $m > 2$ there are more than one orbit type, and with the induced differential structure they are actually not differentiable manifolds. However, this artificial situation can be remedied by "relaxing" the differential structure so that the quotient sphere becomes the standard sphere while the map π is still differentiable, see Section 3.4.3.*

Remark 3.11 *The above orbital fibrations are natural maps between spheres with specific properties, with possible future applications in physics. Indeed, the Hopf fibration (112) is a simple special case with several well known applications. For example, it describes the geometry of a magnetic monopole, and Dirac made the major discovery that the fibration could explain (in a modern language) the quantization of electric charge. Here S^3 is the unit sphere of $\mathbb{R}^4 = \mathbb{C}^2$, the action of $SO(2) \simeq U(1)$ extends to \mathbb{C}^2 by complex scalar multiplication, and the point magnetic source is at the origin of $\mathbb{C}^2/U(1) \approx \mathbb{R}^3$. The 4-space \mathbb{R}^4 in this setting is also referred to as the Kaluza-Klein model of the Dirac monopole.*

Arnold[3] has considered the three special cases $m = 2, 3, 5$ of (109) from a different viewpoint. Namely, they fit into the following unifying pattern with projective spaces over \mathbb{R}, \mathbb{C} or \mathbb{H} as intermediate quotient spaces :

$$\begin{aligned} i) \quad & S^1 \rightarrow \mathbb{R}P^1 \simeq S^1 \\ ii) \quad & S^5 \rightarrow \mathbb{C}P^2 \rightarrow \mathbb{C}P^2/O(1) \approx S^4 \\ iii) \quad & S^{19} \rightarrow \mathbb{H}P^4 \rightarrow \mathbb{H}P^4/\langle Aut\mathbb{H}, \mathbb{Z}_2 \rangle \approx S^{13} \end{aligned} \quad (113)$$

where $O(1) \simeq \mathbb{Z}_2$ acts by (complex) conjugation and $Aut\mathbb{H} \simeq SO(3)$ is the automorphism group of \mathbb{H} (quaternions). The last two cases are two-step orbital fibrations defined by groups $H \rightarrow K \rightarrow K/H$, that is, K acts on the sphere on the left side in (113), and the intermediate (projective) space is the orbit space of the normal subgroup H , with the induced action of K/H . The groups corresponding to the last two cases of (113) are as follows

$$ii) \quad SO(2) \rightarrow O(2) \rightarrow O(1), \quad iii) \quad Sp(1) \rightarrow O(4) \rightarrow O(3)$$

where $O(3) = SO(3) \times \mathbb{Z}_2 \simeq \langle Aut\mathbb{H}, \mathbb{Z}_2 \rangle$. However, from this viewpoint it is also natural to include the cases $m = 2, 4$ of (111), where there is no group \mathbb{Z}_2 . Case $m = 2$ is the Hopf fibration (112) and the new case $m = 4$ reads

$$S^{15} \rightarrow \mathbb{H} P^3 \rightarrow \mathbb{H} P^3 / Aut\mathbb{H} \approx S^9 \quad (114)$$

with the corresponding groups $Sp(1) \rightarrow SO(4) \rightarrow SO(3)$.

3.4.3 The m-universal linear model

We shall describe another property of the m-universal shape space $M^*(m, m) \simeq Sym^*(m)$, which by (104) is known to be a disk whose boundary sphere is $M^*(m - 1, m)$. In fact, what is remarkable is that the equivariant system $(O(m), Sym^*(m))$ has a *linear model*, namely it "resembles" closely a Euclidean disk with $O(m)$ acting as an orthogonal transformation group. To make this correspondence precise, let us first inquire what is the appropriate linear model.

Recall from Section 3.2 that $O(m)$ acts orthogonally on the space $Sym(m)$ of symmetric matrices of dimension m , namely by the symmetric tensor product representation $S^2\rho_m$. Let $Sym^0(m)$ be the subspace of matrices of trace zero, where $O(m)$ acts by the irreducible representation $S^2\rho_m - 1$, and let

$$D^0(m) \subset Sym^0(m) \quad \text{cf. (70)}$$

be the unit disk centered at the origin. Then we make the following definition which will be justified below.

Definition 3.12 *The above equivariant system $(O(m), D^0(m))$ is the linear model of $(O(m), M^*(m, m))$, and the restriction to the boundary sphere $\partial D^0(m)$ is the linear model of $\partial M^*(m, m) = M^*(m - 1, m)$.*

Now, we shall construct a 1-1 correspondence between $Sym^*(m)$ and the disk $D^0(m)$ which is smooth and $O(m)$ -equivariant, that is, the map commutes with the action of $O(m)$. To each matrix $Y \in Sym^*(m)$, namely a positive semidefinite matrix with trace 1, we associate the matrix

$$Y_0 = \frac{1}{r_m} \left(Y - \frac{1}{m} Id \right) \in Sym^0(m), \quad \text{where } r_m = \sqrt{1 - \frac{1}{m}},$$

and note that Y_0 is perpendicular to the identity Id and

$$|Y_0|^2 = \frac{1}{r_m^2} \left(|Y|^2 - \frac{1}{m} \right) = \frac{1}{r_m^2} \left(\text{trace}(Y^2) - \frac{1}{m} \right) \leq \frac{1}{r_m^2} \left(\text{trace}(Y)^2 - \frac{1}{m} \right) = 1$$

Therefore, the $O(m)$ -equivariant affine transformation

$$Sym(m) \rightarrow Sym(m) : Y \rightarrow \frac{1}{r_m} \left(Y - \frac{1}{m} Id \right) = Y_0$$

restricts to an embedding

$$Sym^*(m) \rightarrow D^0(m), \quad Y \rightarrow Y_0 \quad (115)$$

between disks of the same dimension, which maps the *geometric center* $\frac{1}{m}Id$ to the origin, i.e., the center of $D^0(m)$.

Thus, by a translation and homothety inside the Euclidean space $Sym(m)$ the convex subset $Sym^*(m)$ becomes, somehow, an "inward" equivariant deformation of $D^0(m)$, which by (73) fixes the subset $Sym^*(m)_1 \simeq \mathbb{R}P^{m-1}$ lying on the boundary sphere $\partial D^0(m)$. In the simplest case $m = 2$ we have more explicitly

$$Sym^*(2) = \left\{ \begin{pmatrix} u & v \\ v & 1-u \end{pmatrix} ; (u - \frac{1}{2})^2 + v^2 \leq \frac{1}{4} \right\} \simeq D^2 \quad (116)$$

and the embedding (115) is actually a diffeomorphism.

However, the embedding (115) is not surjective for $m > 2$, so let us explain how to further deform equivariantly to make the embedding fill the whole unit disk. First of all, by the convexity of $Sym^*(m)$ it follows that each ray in $D^0(m)$ from the origin passes through a unique point of the embedded sphere $\partial Sym^*(m)$. In particular, by an additional scaling we obtain a 1-1 correspondence between the boundary spheres of the two disks, and the final composition

$$\psi_m : \partial Sym^*(m) \rightarrow \partial D^0(m) = S^{m(m+1)/2-2}, \quad Y \rightarrow Y_0 \rightarrow \frac{1}{|Y_0|} Y_0 \quad (117)$$

is certainly an $O(m)$ -equivariant and smooth homeomorphism.

On the other hand, the above map may be extended to the whole disk, as follows. First of all, each ray from the origin intersects the embedded disk $Sym^*(m)$ in a segment. So, let us stretch the segment outward along the ray until it has unit length, but with no stretching in a neighborhood of the origin. Moreover, the stretching must be specified by a function on the orbit space since $O(m)$ -equivalent segments must be stretched in the same way to make ψ_m equivariant. Following this procedure we certainly obtain an $O(m)$ -equivariant and smooth homeomorphism

$$\psi_m : Sym^*(m) \rightarrow \approx D^0(m) \quad (118)$$

which extends the map in (117). Such an equivariant and smooth homeomorphism is not unique, of course. The ambiguity lies in the group $Diff_{O(m)}(D^0(m))$ consisting of all equivariant diffeomorphisms of the linear model. Yet, another construction of equivariant homeomorphisms like ψ_m is described briefly at the end of this subsection.

In summary, we arrive at the following result :

Theorem 3.13 *There is an $O(m)$ -equivariant and differentiable homeomorphism ψ_m from the m -universal shape space $M^*(m, m) \simeq Sym^*(m)$ to its linear model $D^0(m) \simeq D^{m(m+1)/2-1}$. For $m = 2$ this map is a diffeomorphism.*

By referring to the following diagram

$$\begin{array}{ccccc} & & M^1(m, m) & & \\ & & \downarrow \pi & & \\ \mathbb{R} & \xleftarrow{g} & Sym^*(m) & \xrightarrow{\psi_m} & D^0(m) \end{array} \quad (119)$$

let us explain the differential structure of $Sym^*(m)$ induced via the orbit map π . Namely, the function g is said to be smooth if the composed map $g \circ \pi$ is smooth. Moreover, smoothness of ψ_m means the composed map $\psi_m \circ \pi$ is smooth, and therefore we actually know (by our construction) that ψ_m is smooth.

On the other hand, for $m > 2$ there also exists a smooth map g such that $g \circ \psi_m^{-1}$ is not smooth, as a function on the Euclidean disk $D^0(m)$, and consequently ψ_m^{-1} cannot be smooth. Here we also refer to the discussion at the end of Section 3.4.4, together with Lemma 3.9 and Remark 3.10. Briefly, the differential structure on $Sym^*(m)$ induced via π is a strict "refinement" of the *standard* structure which $Sym^*(m)$ inherits from the Euclidean disk (via ψ_m).

Remark 3.14 *We have seen that the m -universal shape space $Sym^*(m)$ has two naturally induced smooth structures, namely induced via π and ψ_m^{-1} respectively, and they are different when $m > 2$. In fact, $Sym^*(m)$ is not a smooth manifold in the first case. But in both cases π and ψ_m are smooth maps, but the standard structure (induced via ψ_m^{-1}) is the only one that makes ψ_m an $O(m)$ -equivariant diffeomorphism.*

Finally, let us turn to the natural simplicial structures of the orbit spaces of $Sym^*(m)$ and the disk $D^0(m)$, namely the following two spherical simplices

$$\begin{aligned} \frac{Sym^*(m)}{O(m)} &\simeq \Delta^{m-1} = \frac{S^{m-1}}{B_m} : r_1 \geq r_2 \geq \dots \geq r_m \geq 0, \sum r_i^2 = 1 \quad (120) \\ \frac{D^0(m)}{O(m)} &\simeq \bar{\Delta}^{m-1} = \frac{D^{m-1}}{S_m} : \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m, \sum \lambda_i = 0, \sum \lambda_i^2 \leq 1 \end{aligned}$$

Despite the above theorem, which implies that ψ_m induces an $O(m)$ -orbit type strata preserving homeomorphism $\bar{\psi}_m : \Delta^{m-1} \rightarrow \bar{\Delta}^{m-1}$, the two structures in (120) are conspicuously different. The reason is that the simplicial structure in the first case reflects the induced subrank stratification of $Sym^*(m)$ (cf. Section 3.3.3), namely the common refinement of the rank and $O(m)$ -orbit type stratification, whereas the simplicial structure in the second case is merely reflecting the pure $O(m)$ -orbital stratification of the orthogonal transformation group $O(m)$ on $D^0(m)$. Thus, by passing from $Sym^*(m)$ to its linear model $D^0(m)$ the notion of rank is seemingly lost. Therefore, let us also investigate how the rank strata can be recognized in the linear model itself.

Recall that $M^*(k, m) \subset Sym^*(m)$ is the subspace lying above $\Delta^{k-1} : r_{k+1} = \dots = r_m = 0$, and the rank k stratum is defined by the subset $\Delta_k^{k-1} \subset \Delta^{k-1} : r_k > 0$. This stratum has all the $O(m)$ -orbit types

$$O(k_1, k_2, \dots, k_p) \times O(m - k), \quad \sum k_i = k, \text{ cf. (89)} \quad (121)$$

labelled by the various subranks $\kappa = (k_1, \dots, k_p)$ which record the strings of equalities among the numbers $r_i > 0$, see (86), (87).

On the other hand, in the linear disk model any pattern of equality strings among the λ_i 's corresponds in the same way to a tuple $\kappa' = (k_1, \dots, k_p, k_{p+1})$ of positive integers, where k_1 is the number of λ_i 's in (120) equal to λ_1 , k_2 is the number of λ_i 's equal to λ_{k_1+1} etc., and the factor $O(m - k)$ in (121) is

replaced by $O(k_{p+1})$. In any case, the last factor $O(m - k)$ (resp. $O(k_{p+1})$) of the isotropy group (121) is no more special than the other factors $O(k_i)$, and this clearly explains why the rank is not determined by the $O(m)$ -orbit type.

We claim, however, the rank k is determined in the linear model by the identity $k_{p+1} = m - k$ provided the obvious condition $\sum \lambda_i^2 = 1$ holds. The latter condition merely says $M^*(k, m)$ embeds as a subset of the sphere $\partial D^0(m)$ and hence away from the interior $D^0(m)$ of the disk. The interior is, of course, given by $\sum \lambda_i^2 < 1$ and here the rank is $k = m$. Moreover, by removing k_{p+1} from the tuple κ' we are left with the correct subrank tuple κ . The above claim about k will be settled below.

Any $O(m)$ -equivariant homeomorphism

$$\psi : \text{Sym}^*(m) \rightarrow^\approx D^0(m)$$

such as ψ_m , for example, induces an $O(m)$ -orbital strata preserving homeomorphism $\bar{\psi} : \Delta^{m-1} \rightarrow \bar{\Delta}^{m-1}$. Conversely, let us see how to start from $\bar{\psi}$ and construct an appropriate lifting ψ as above. The idea is to construct $\bar{\psi}$ as a map between fundamental domains with the $O(m)$ -isovariant property, that is, a point and its image point have the same isotropy group. Then ψ will be the unique equivariant extension to all of $\text{Sym}^*(m)$. To choose appropriate fundamental domains, first observe that $\bar{\Delta}^{m-1} \subset D^0(m)$ is a fundamental domain if the numbers λ_i in (120) are regarded as the entries of a diagonal matrix. Similarly, the following subset of diagonal matrices

$$\tilde{\Delta}^{m-1} \subset \text{Sym}^*(m) : \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq 0, \sum \lambda_i = 1$$

is a fundamental domain in $\text{Sym}^*(m)$. Then the $O(m)$ -orbit map projection $\tilde{\Delta}^{m-1} \rightarrow \Delta^{m-1}$ is just $\lambda_i \rightarrow \sqrt{\lambda_i} = r_i$, see (83), and we note that $\partial \tilde{\Delta}^{m-1} \simeq \tilde{\Delta}^{m-2}$ is the subset with $\lambda_m = 0$.

For example, the above equivariant homeomorphism (118) between the boundary spheres corresponds to the $O(m)$ -isovariant map

$$\frac{S^{m-2}}{B_{m-1}} \simeq \tilde{\Delta}^{m-2} = \partial \tilde{\Delta}^{m-1} \xrightarrow{\bar{\psi}_m} \partial \bar{\Delta}^{m-1} \simeq \frac{S^{m-2}}{S_m} \quad (122)$$

which sends $Y = \text{diag}(\lambda_1, \dots, \lambda_{m-1}, 0)$ to the matrix

$$\frac{1}{|Y_0|} Y_0 = \frac{1}{\lambda} \text{diag}(\lambda_1 - 1/m, \dots, \lambda_{m-1} - 1/m, -1/m) \quad (123)$$

where $\lambda = \sqrt{\lambda_1^2 + \dots + \lambda_{m-1}^2 - 1/m}$.

The map (122) is isovariant since the inequality pattern among the entries of the matrix Y and its image is preserved. From this it is also clear that $\lambda_{k+1} = 0$ if and only if the $m - k$ last entries in (123) are identical. This also settles the above claim concerning the rank k recognition in the linear model.

3.4.4 The simplest shape spaces

We return to the shape spaces $M^*(3, m)$ together with their $SO(3)$ -version for $m = 2, 3, 4$. Recall from Section 1 that $M^*(3, n - 1) = M_n^*$ is the shape space

for the n -body problem, and for $n = 3, 4, 5$ the topological classification of these spaces are as follows :

$$\begin{aligned}
n = 3 : M_3^* &\simeq \frac{M^1(3, 2)}{O(3)} = \frac{M^1(3, 2)}{SO(3)} = \frac{M^1(2, 2)}{O(2)} = D^2 \\
n = 4 : M_4^* &\simeq \frac{M^1(3, 3)}{O(3)} \approx D^5, \quad \frac{M^1(3, 3)}{SO(3)} \approx 2D^5 = S^5 \\
n = 5 : M_5^* &\simeq \frac{M^1(3, 4)}{O(3)} \approx S^8, \quad \frac{M^1(3, 4)}{SO(3)} \approx 2S^8 = S^8 \cup_P S^8 \quad (P = \mathbb{C}P^3/\mathbb{Z}_2)
\end{aligned} \tag{124}$$

We shall make some further comments on the cases $n = 3, 4$.

$n = 3$: With the induced metric the disk D^2 is a hemisphere of the base space

$$2D^2 = \frac{M^1(2, 2)}{SO(2)} = \mathbb{C}P^1 = S^2(1/2) \tag{125}$$

of the Hopf fibration (112), which is a round sphere of radius $1/2$. Its equator circle $M^*(1, 2) = \mathbb{R}P^1 = S^1(1/2)$ represents the shapes of degenerate (i.e. collinear) 3-configurations. But in the study of the 3-body problem it is, in fact, natural to use the whole sphere as the shape space. The reason is that a non-degenerate 3-configuration in 3-space is geometrically a triangle which can be oriented in two different ways, by the ordering of the vertices. Then the two hemispheres in (125) represent triangles with opposite orientation, cf. [10], [12]. In this way a 3-body motion corresponds to a continuous motion of a mass triangle whose orientation changes when the motion passes through an eclipse, that is, when the shape curve crosses the equator circle.

$n = 4$: The boundary of the disk D^5 in (124) is the shape space of coplanar 4-configurations, that is, the sphere

$$M^*(2, 3) = \mathbb{C}P^2/\mathbb{Z}_2 \approx S^4 \quad (\text{cf. (109) with } m = 3) \tag{126}$$

and the action of its isometry group $O(3)$ is equivalent (by Theorem 3.13) to that of its linear model $(O(3), S^4, S^2\rho_3 - 1)$. Namely, the linear model is the space of symmetric 3×3 -matrices with zero trace and unit norm, with the natural action of $O(3)$ by conjugation.

The fact that the quotient space $\mathbb{C}P^2/\mathbb{Z}_2$ is homeomorphic to S^4 was already known to L.S. Pontryagin in the 1930's, according to Arnold[3], and we refer to [2], [15], [21] for different proofs of this specific result. Massey also observed that the induced $SO(3)$ -action on $\mathbb{C}P^2/\mathbb{Z}_2$ has the same orbit structure as that of the above linear model $(SO(3), S^4)$, and the existence and construction of an equivariant homeomorphism (such as (117)) was, in fact, a problem stated by Massey. Moreover, Arnold[2] has constructions which are very close to our linear model construction.

Arnold also discusses the differentiable structure of S^4 , as a quotient space $\mathbb{C}P^2/\mathbb{Z}_2$, and it is stated that $\mathbb{C}P^2/\mathbb{Z}_2$ is diffeomorphic to S^4 (cf. §1 in [2]). However, what is shown is that the composite map

$$M^1(2, 3) = S^5 \rightarrow \mathbb{C}P^2 \rightarrow \mathbb{C}P^2/\mathbb{Z}_2 \rightarrow^\approx S^4 \quad (127)$$

from S^5 to S^4 is differentiable (in the usual sense). Both $\mathbb{C}P^2$ and $\mathbb{C}P^2/\mathbb{Z}_2$ have the induced smooth functional structure as quotient spaces of S^5 , but only $\mathbb{C}P^2$ becomes a differentiable manifold in this way. So, Arnold "relaxes" the differential structure on $\mathbb{C}P^2/\mathbb{Z}_2$ so that the last map in (127) becomes a diffeomorphism, that is, he defines the differential structure to be the *standard* structure mentioned in Remark 3.14. This makes the last map in (127) a diffeomorphism while the composed map $S^5 \rightarrow S^4$ in (127) is still a smooth map. The same applies, of course, to all the constructions $S^p \rightarrow S^p/K \rightarrow^\approx S^q$ in Section 3.4.2, where none of the "quotient spheres" S^p/K is really a smooth manifold when $q \geq 4$.

Finally, let us explain why the "quotient spheres" are not smooth manifolds. Since all cases are analogous we consider again the simplest case

$$M^*(2, 3) = S^5/O(2) = \mathbb{C}P^2/\mathbb{Z}_2 \approx S^4 \quad (128)$$

viewed as an $O(2)$ -orbit space with the induced smooth structure. Then our claim is that the orbit space in (128) is not a differentiable manifold in a neighborhood of the subset $A = M^*(1, 3) \simeq \mathbb{R}P^2$. To see this we shall apply the slice theorem (cf. Section 3.1.1), according to which each $O(2)$ -orbit in S^5 belongs to a tubular neighborhood U whose image $\bar{U} = U/O(2)$ in the orbit space is an open set of type (57), namely diffeomorphic to the orbit space of the slice representation. The set A represents those orbits of type $O(1)$, and the slice representation of $O(1)$ acts on \mathbb{R}^4 with the eigenvalues $(-1, -1, 1, 1)$, consequently any point on A has an open neighborhood diffeomorphic to

$$\bar{U} = \frac{O(2) \times_{O(1)} (\mathbb{R}^2 \oplus \mathbb{R}^2)}{O(2)} \simeq \frac{\mathbb{R}^2 \oplus \mathbb{R}^2}{O(1)} \simeq \frac{\mathbb{R}^2}{O(1)} \times \mathbb{R}^2 \quad (129)$$

where the first factor is of type (102) and is transversal to the set A and the second is $\simeq \bar{U} \cap A$. However, the product space in (129) is homeomorphic to \mathbb{R}^4 , but in view of Lemma 3.9 it is not a differentiable manifold.

4 Geometric invariants of n-body systems

By a *geometric invariant* on the centered configuration space

$$M_n : \sum m_i \mathbf{a}_i = \mathbf{0}$$

we mean a polynomial function $F(\mathbf{a}_1, \dots, \mathbf{a}_n)$ which is invariant under congruence and internal symmetries. We will describe the ring of all these invariants by calculating the ring of invariants for matrix spaces $M(d, m)$ in general. Denote a typical matrix by

$$X = (\mathbf{x}_1, \dots, \mathbf{x}_m) = [\mathbf{x}_1^*, \dots, \mathbf{x}_d^*]$$

where the \mathbf{x}_i and \mathbf{x}_j^* are the column and row vectors respectively. The product group $GL(d) \times GL(m)$ acts on $M(d, m)$ by matrix multiplication

$$X \rightarrow \psi X \varphi^t, \quad (\psi, \varphi) \in GL(d) \times GL(m),$$

and $O(d) \times O(m)$ is the subgroup leaving invariant the standard metric form

$$I = |X|^2 = \text{trace}(X^t X) = \sum_{i=1}^m |\mathbf{x}_i|^2 = \sum_{j=1}^d |\mathbf{x}_j^*|^2, \quad (130)$$

On the other hand, for fixed mass distribution $\mu = (m_1, \dots, m_m)$, $m_i > 0$, let us also consider the mass dependent metric defined by

$$I^{(\mu)} = \text{trace}(DX^t XD) = \sum_{i=1}^m m_i |\mathbf{x}_i|^2, \quad (131)$$

where $D = \text{diag}(\sqrt{m_1}, \sqrt{m_2}, \dots, \sqrt{m_m})$ and the corresponding isometry subgroup is $O(d) \times O^*(m) \subset GL(d) \times GL(m)$, where

$$O^*(m) = DO(m)D^{-1} \subset GL(m) \quad (132)$$

is the subgroup leaving invariant the metric form $\sum m_i x_i^2$ on \mathbb{R}^m (space of row vectors with right side action of $GL(m)$).

Lemma 4.1 *The invariant ring on $M = M(d, m)$, $d \leq m$, under the action of $O(d) \times O^*(m)$, is the polynomial ring with generators*

$$I_k = \sum_{i_1 < i_2 < \dots < i_k} m_{i_1} m_{i_2} \dots m_{i_k} \cdot |\mathbf{x}_{i_1} \wedge \mathbf{x}_{i_2} \wedge \dots \wedge \mathbf{x}_{i_k}|^2, \quad 1 \leq k \leq d, \quad (133)$$

where $I_1 = I^{(\mu)}$ and the exterior product space $\wedge^k \mathbb{R}^m$ has the standard norm.

Proof. First assume $m_i = 1$ for all i . It is easy to verify the identity

$$I_k = \sum_{i_1 < i_2 < \dots < i_k} |\mathbf{x}_{i_1} \wedge \mathbf{x}_{i_2} \wedge \dots \wedge \mathbf{x}_{i_k}|^2 = \sum_{i_1 < i_2 < \dots < i_k} |\mathbf{x}_{i_1}^* \wedge \mathbf{x}_{i_2}^* \wedge \dots \wedge \mathbf{x}_{i_k}^*|^2 \quad (134)$$

for each $k = 1, \dots, d$. Then, from the column vector version it is clear that I_k is invariant under the action of $O(d)$, and similarly the row vector expression is invariant under the action of $O(m)$. Hence, I_k is an invariant of G .

Let us apply the *reduction principle* for orthogonal transformation groups (cf. e.g. [28]). Namely, we first calculate the *reduced group* of the action of $G = O(d) \times O(m)$ on M , which is the quotient group

$$\bar{G} = N_G(H)/H \simeq B_d$$

and is finite in our case, where

$$H = \Delta O(1)^d \times O(m-d) \subset G, \quad \text{cf. (89)}$$

is the principal isotropy group and $N_G(H)$ is its normalizer in G . Next we determine the fixed point set of H

$$\mathbb{R}^d = F(H, M) : X = \text{diag}(r_1, r_2, \dots, r_d), \quad \text{cf. (81)}$$

which contains the fundamental domain

$$M^+(d) \subset \mathbb{R}^d \subset M : r_1 \geq r_2 \geq \dots \geq r_d \geq 0$$

of the G -action, identified with the orbit space of G in Section 3.3.1. Moreover, for the G -orbit through X , $r_i^2 = \lambda_i$ are the eigenvalues of $X^t X$.

Finally, we consider the induced action of \bar{G} on \mathbb{R}^d , which is a group generated by reflections and as such it is the Weyl group B_d , having the above cone $M^+(d)$ as a fundamental domain. Now, the reduction principle says that the restriction of polynomials induces an isomorphism between the invariant rings of (G, M) and (\bar{G}, \mathbb{R}^d) . But the latter ring is certainly generated by the elementary symmetric functions

$$I_1 = \sum \lambda_i, \quad I_2 = \sum_{i < j} \lambda_i \lambda_j, \quad I_3 = \sum_{i < j < k} \lambda_i \lambda_j \lambda_k, \quad \dots$$

which are also the restrictions of the functions in (134), and this proves the lemma for masses $m_i = 1$.

Next, let $M(d, m)^{(\mu)}$ be the matrix space with the mass dependent metric (131) and consider the transformation

$$M(d, m)^{(\mu)} \rightarrow M(d, m), \quad X \rightarrow XD = (\sqrt{m_1} \mathbf{x}_1, \sqrt{m_2} \mathbf{x}_2, \dots, \sqrt{m_m} \mathbf{x}_m)$$

where the left and right hand space have metrics (131) and (130) respectively. The transformation is a G -equivariant isometry, where $O(m)$ acts on $M(d, m)^{(\mu)}$ by first applying the isomorphism

$$O(m) \rightarrow DO(m)D^{-1} = O^*(m), \quad \text{cf. (132)}$$

Hence, the transformation also induces an isomorphism of the corresponding invariant rings. In effect, the invariant ring of $M(d, m)^{(\mu)}$ is obtained by replacing each vector \mathbf{x}_i in (134) by $\sqrt{m_i} \mathbf{x}_i$. ■

In particular, the centered configuration space $M_n \simeq M(3, n-1)$ has the following basic geometric invariants, in terms of the Jacobi vectors and classical vector operations :

$$I_1 = \sum |\mathbf{x}_i|^2, \quad I_2 = \sum_{i < j} |\mathbf{x}_i \times \mathbf{x}_j|^2, \quad I_3 = \sum_{i < j < k} |(\mathbf{x}_i \times \mathbf{x}_j) \cdot \mathbf{x}_k|^2 \quad (135)$$

for all $n \geq 4$. (For $n = 3$ the generators are I_1 and I_2). The expressions for I_k hold for any choice of Jacobi vectors, but still it remains to express them as functions of $\mathbf{a}_1, \dots, \mathbf{a}_n$. To this end, consider the free n -configuration space or matrix space

$$\hat{M}_n = M(3, n)^{(\mu)} = \{X = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)\}$$

with the metric form $I^{(\mu)} = \sum m_i |\mathbf{a}_i|^2$ as in (131). By the above lemma its invariant ring with respect to the group $O(3) \times O^*(n)$ has the basic generators

$$\begin{aligned} I_1 &= \sum m_i |\mathbf{a}_i|^2, \quad I_2 = \sum_{i < j} m_i m_j |\mathbf{a}_i \times \mathbf{a}_j|^2, \\ I_3 &= \sum_{i < j < k} m_i m_j m_k |(\mathbf{a}_i \times \mathbf{a}_j) \cdot \mathbf{a}_k|^2 \end{aligned} \quad (136)$$

and these are still complete and independent as invariants of the subgroup $O(3) \times O^*(n-1)$ acting on the subspace $M_n \subset \hat{M}_n$, as long as $3 \leq n-1$. Here $O^*(n-1) \subset O^*(n)$ is the internal symmetry group of M_n and is the subgroup leaving M_n invariant, or equivalently, the subgroup of $O^*(n)$ which under the action on \mathbb{R}^n (row vectors) fixes the vector (m_1, m_2, \dots, m_n) . This proves the following result.

Theorem 4.2 *The ring of geometric invariants of the centered n -configuration space M_n is the polynomial ring generated by I_1, \dots, I_q in (136), $q = \min\{3, n-1\}$.*

Remark 4.3 *Starting from the Jacobi vector expressions (135) of the invariants, a Jacobi transformation may transform them to expressions involving the vectors \mathbf{a}_i . In this way, however, we may arrive at a non-symmetric expression since $\sum m_i \mathbf{a}_i = \mathbf{0}$. Below we shall give examples to illustrate the non-uniqueness of the symmetrization procedure and also give a geometric interpretation of I_{n-1} for $n = 3$ or 4.*

- $n = 3$: Let \mathcal{A} be the area of the triangle spanned by the vector triple $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ with $\sum m_i \mathbf{a}_i = \mathbf{0}$. By simple trigonometry

$$2\mathcal{A} = \frac{\bar{m}}{m_3} |\mathbf{a}_1 \times \mathbf{a}_2| = \frac{\bar{m}}{m_1} |\mathbf{a}_2 \times \mathbf{a}_3| = \frac{\bar{m}}{m_2} |\mathbf{a}_3 \times \mathbf{a}_1|$$

On the other hand, from the Jacobi vector formula (43),

$$\mathbf{x}_1 \times \mathbf{x}_2 = \zeta_1 \zeta_2 \mathbf{a}_1 \times \mathbf{a}_2$$

and consequently

$$|\mathbf{x}_1 \times \mathbf{x}_2| = \zeta_1 \zeta_2 |\mathbf{a}_1 \times \mathbf{a}_2| = 2 \sqrt{\frac{m_1 m_2 m_3}{\bar{m}}} \mathcal{A}$$

and

$$\begin{aligned} I_2 &= |\mathbf{x}_1 \times \mathbf{x}_2|^2 = \bar{m} \frac{m_i m_j}{m_k} |\mathbf{a}_i \times \mathbf{a}_j|^2 \\ &= \bar{m} \frac{m_1 m_2}{3m_3} |\mathbf{a}_1 \times \mathbf{a}_2|^2 + \bar{m} \frac{m_2 m_3}{3m_1} |\mathbf{a}_2 \times \mathbf{a}_3|^2 + \bar{m} \frac{m_3 m_1}{3m_2} |\mathbf{a}_3 \times \mathbf{a}_1|^2 \end{aligned}$$

where $\{i, j, k\} = \{1, 2, 3\}$ and the last sum is a symmetrization with all three terms equal, in fact. However, the above expression for I_2 and the symmetric expression in (136) are identical as a function on M_3 .

- $n = 4$: Let \mathcal{V} be the volume of the tetrahedron spanned by the vector quadruple $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4)$ satisfying $\sum m_i \mathbf{a}_i = 0$. Then $\pm 6\mathcal{V}$ equals the triple product of $\mathbf{a}_i - \mathbf{a}_1$, $i = 2, 3, 4$, and

$$6\mathcal{V} = \frac{\bar{m}}{m_4} |(\mathbf{a}_1 \times \mathbf{a}_2) \cdot \mathbf{a}_3| = \frac{\bar{m}}{m_1} |(\mathbf{a}_2 \times \mathbf{a}_3) \cdot \mathbf{a}_4| \text{ etc.}$$

Again, by expressing the Jacobi vectors \mathbf{x}_i in terms of the \mathbf{a}_i it follows

$$|(\mathbf{x}_1 \times \mathbf{x}_2) \cdot \mathbf{x}_3| = \zeta_1 \zeta_2 \zeta_3 |(\mathbf{a}_1 \times \mathbf{a}_2) \cdot \mathbf{a}_3| = 3! \sqrt{\frac{m_1 m_2 m_3 m_4}{\bar{m}}} \mathcal{V}$$

and consequently

$$\begin{aligned} I_3 &= |(\mathbf{x}_1 \times \mathbf{x}_2) \cdot \mathbf{x}_3|^2 = \bar{m} \frac{m_i m_j m_k}{m_l} |(\mathbf{a}_i \times \mathbf{a}_j) \cdot \mathbf{a}_k|^2 \\ &= \bar{m} \frac{m_1 m_2 m_3}{4m_4} |(\mathbf{a}_1 \times \mathbf{a}_2) \cdot \mathbf{a}_3|^2 + \bar{m} \frac{m_2 m_3 m_4}{4m_1} |(\mathbf{a}_2 \times \mathbf{a}_3) \cdot \mathbf{a}_4|^2 + \text{etc.} \end{aligned}$$

where $\{i, j, k, l\} = \{1, 2, 3, 4\}$ and the four terms in the last sum are equal. Again, I_3 in (136) is an another symmetric expression for the same function on M_4 .

Remark 4.4 *The cases $n = 3$ or 4 are special since, for example, their shape space M_n^* is topologically a disk D^2 or D^5 respectively. Moreover, in these cases M_n^* has a unique geometric center, and this may be characterized in several ways. It is the unique fixed point of the symmetry group $O(n-1)$, and the point has the same distance to each of the binary collision varieties. Furthermore, it is also the maximum point of I_2 or I_3 respectively. That is, for a fixed size $I_1 = 1$ the triangle (resp. tetrahedron) with the largest area (resp. volume) is at the geometric center. In terms of Jacobi vectors the condition is that $\mathbf{x}_i \cdot \mathbf{x}_j = 0$ and $|\mathbf{x}_i| = |\mathbf{x}_j|$ for $i \neq j$.*

5 The weighted root system of an n-body system

A typical potential function $U(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ of an n-body problem depends only on the pairwise distances

$$r_{ij} = |\mathbf{a}_i - \mathbf{a}_j|, \quad 1 \leq i < j \leq n,$$

and for certain applications one would like to express them in terms of Jacobi vectors \mathbf{x}_i , say

$$X = (\mathbf{x}_1, \dots, \mathbf{x}_{n-1}) \in M(3, n-1)$$

is the matrix associated with a given Jacobi transformation $\Psi : M_n \simeq M(3, n-1)$. For this purpose we shall define and investigate the *weighted root system*, depending on the mass distribution, which is primarily a book-keeping for expanding the vectors $\mathbf{a}_i - \mathbf{a}_j$ as a linear combination of Jacobi vectors. It is naturally a "weighted" version of the root system of type A_{n-1} in classical Lie theory, and they are identical when the mass distribution is uniform.

5.1 Distance functions and the Ψ -root system

The mutual distances r_{ij} have a nice geometric interpretation as functions on M_n which measure the distance from specific subvarieties. To explain this, consider the linear transformation

$$M = M(3, n-1) \rightarrow \mathbb{R}^3$$

defined by

$$X \rightarrow [\mathbf{u}, X] = \sum_{i=1}^{n-1} c_i \mathbf{x}_i = X \mathbf{u}^t \quad (137)$$

where $\mathbf{u} = (c_1, c_2, \dots, c_{n-1}) \in \mathbb{R}^{n-1}$ is a given row vector of unit length and \mathbf{u}^t is the column vector. Define the codimension 3 subspace $B = B_{\mathbf{u}} \subset M$ to be the kernel

$$B : [\mathbf{u}, X] = 0.$$

It is easy to verify that the distance in M from a point X to the subspace B is given by the function

$$\delta_B(X) = \text{dist}(X, B) = |[\mathbf{u}, X]| = \left| \sum_{k=1}^3 (\mathbf{u} \cdot \mathbf{x}_k^*)^2 \right|^{\frac{1}{2}}$$

where the \mathbf{x}_k^* are the row vectors of X and the inner product on \mathbb{R}^{n-1} is the standard one.

In particular, the binary collision varieties

$$B_{ij} \subset M_n : \mathbf{a}_i = \mathbf{a}_j, \quad i \neq j \quad (138)$$

have associated distance functions δ_{ij} of this kind, defined on M via a chosen Jacobi transformation $\Psi : M_n \rightarrow M$. Consequently there is a unit vector (unique up to sign)

$$\mathbf{u}_{ij} = (c_{ij}^1, c_{ij}^2, \dots, c_{ij}^{n-1}) \in \mathbb{R}^{n-1} \quad (139)$$

depending on Ψ and the mass distribution, so that

$$\delta_{ij}(X) = \left| \sum_{k=1}^3 (\mathbf{u}_{ij} \cdot \mathbf{x}_k^*)^2 \right|^{\frac{1}{2}} \quad (140)$$

Let us express δ_{ij} in terms of the n -configuration $\mathbf{X} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$. For \mathbf{X} fixed, let $\mathbf{Y}_0 = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ be the critical point of the distance function $\delta(\mathbf{Y}) = |\mathbf{Y} - \mathbf{X}|$ defined on B_{ij} , where by definition $\mathbf{b}_i = \mathbf{b}_j$. The condition $\nabla \delta = 0$ implies

$$\mathbf{b}_i = \mathbf{b}_j = \frac{m_i \mathbf{a}_i + m_j \mathbf{a}_j}{m_i + m_j}, \quad \mathbf{b}_k = \mathbf{a}_k \text{ for } k \neq i, j$$

and consequently in terms of reduced masses (29)

$$\begin{aligned}\delta_{ij}^2(\mathbf{X}) &= |\mathbf{Y}_0 - \mathbf{X}|^2 = m_i |\mathbf{b}_i - \mathbf{a}_i|^2 + m_j |\mathbf{b}_j - \mathbf{a}_j|^2 \\ &= \frac{m_i m_j}{m_i + m_j} |\mathbf{a}_i - \mathbf{a}_j|^2 = \mu_{ij} r_{ij}^2\end{aligned}$$

that is,

$$r_{ij} = \frac{1}{\sqrt{\mu_{ij}}} \delta_{ij} = \frac{1}{\sqrt{\mu_{ij}}} |[\mathbf{u}_{ij}, X]| \quad (141)$$

Therefore, we may define vectors $\mathbf{u}_{ij} = -\mathbf{u}_{ji}$ uniquely by the constraint

$$\mathbf{a}_i - \mathbf{a}_j = \frac{1}{\sqrt{\mu_{ij}}} [\mathbf{u}_{ij}, X], \quad i \neq j, \quad (142)$$

and clearly they satisfy all identities of type

$$\frac{1}{\sqrt{\mu_{ij}}} \mathbf{u}_{ij} + \frac{1}{\sqrt{\mu_{jk}}} \mathbf{u}_{jk} = \frac{1}{\sqrt{\mu_{ik}}} \mathbf{u}_{ik}. \quad (143)$$

Definition 5.1 *The Ψ -root system with mass distribution (m_1, \dots, m_n) is the above collection of $\binom{n}{2}$ vector pairs $\pm \frac{1}{\sqrt{\mu_{ij}}} \mathbf{u}_{ij} \in \mathbb{R}^{n-1}, 1 \leq i < j \leq n$. The collection $\{\pm \mathbf{u}_{ij}\}$ of unit vectors is the normalized Ψ -root system.*

5.2 The standard weighted root system

It will be useful to have explicit formulas for the Ψ_0 -roots, where

$$\Psi_0 : M_n \rightarrow M(3, n-1)$$

is the standard Jacobi transformation constructed in Section 2.3. The notation used for various constants in Section 2.3 is also used below. We shall also use the notation foBy definition of Ψ_0 (cf. (43), (47))

$$\mathbf{a}_i - \mathbf{a}_j = \frac{\zeta_i}{m_i} \mathbf{x}_i - \frac{1}{\zeta_j} \mathbf{x}_j + \sum_{k=i+1}^{j-1} \frac{\zeta_k}{m^{(k-1)}} \mathbf{x}_k = \frac{1}{\sqrt{\mu_{ij}}} \sum_{k=1}^{n-1} c_{ij}^k \mathbf{x}_k, \quad i < j,$$

(where the term $\frac{1}{\zeta_j} \mathbf{x}_j$ is undefined when $j = n$) and the vector $\mathbf{u}_{ij} = (c_{ij}^1, \dots, c_{ij}^{n-1})$ have its nonzero components c_{ij}^k for k in the range $i \leq k \leq \min\{j, n-1\}$. In particular,

$$\mathbf{u}_{n-1, n} = (0, 0, \dots, 1) \quad (144)$$

and for $i < n-1$

$$\frac{1}{\sqrt{\mu_{i, i+1}}} \mathbf{u}_{i, i+1} = (0, \dots, 0, \sqrt{\frac{m^{(i-1)}}{m_i m^{(i)}}}, -\sqrt{\frac{m^{(i+1)}}{m_{i+1} m^{(i)}}}, 0, \dots, 0) \quad (145)$$

$$\frac{1}{\sqrt{\mu_{in}}} \mathbf{u}_{in} = (0, \dots, 0, \sqrt{\frac{m^{(i-1)}}{m_i m^{(i)}}}, \sqrt{\frac{m_{i+1}}{m^{(i)} m^{(i+1)}}}, \dots, \sqrt{\frac{m_{n-1}}{m^{(i)} m^{(n-1)}}}) \quad (146)$$

whereas for $2 \leq i+1 < j < n$

$$\begin{aligned} \frac{1}{\sqrt{\mu_{ij}}} \mathbf{u}_{ij} = & (0, \dots, 0, \sqrt{\frac{m^{(i-1)}}{m_i m^{(i)}}}, \sqrt{\frac{m_{i+1}}{m^{(i)} m^{(i+1)}}}, \dots, \sqrt{\frac{m_k}{m^{(k-1)} m^{(k)}}}, \dots \\ & , \sqrt{\frac{m_{j-1}}{m^{(j-2)} m^{(j-1)}}}, -\sqrt{\frac{m^{(j)}}{m_j m^{(j-1)}}}, 0, \dots, 0) \end{aligned} \quad (147)$$

For a fixed mass distribution, all Ψ -root systems are, in fact, orthogonally equivalent to the above Ψ_0 -root system $\left\{ \pm \frac{1}{\sqrt{\mu_{ij}}} \mathbf{u}_{ij}, i < j \right\}$, see the following subsection.

Example 5.2 $n = 3$:

$$\begin{aligned} \mathbf{u}_{12} = & \left(\sqrt{\frac{m_2 \bar{m}}{(m_2 + m_3)(m_1 + m_2)}}, -\sqrt{\frac{m_1 m_3}{(m_2 + m_3)(m_1 + m_2)}} \right) \\ \mathbf{u}_{13} = & \left(\sqrt{\frac{m_3 \bar{m}}{(m_2 + m_3)(m_1 + m_2)}}, \sqrt{\frac{m_1 m_2}{(m_2 + m_3)(m_1 + m_2)}} \right) \end{aligned}$$

Example 5.3 $n \geq 3$:

$$\begin{aligned} n \geq 3 : \frac{1}{\sqrt{\mu_{12}}} \mathbf{u}_{12} = & \left(\sqrt{\frac{\bar{m}}{m_1(m_2 + \dots + m_n)}}, -\sqrt{\frac{m_3 + \dots + m_n}{m_2(m_2 + \dots + m_n)}}, 0, \dots \right) \\ n \geq 4 : \frac{1}{\sqrt{\mu_{23}}} \mathbf{u}_{23} = & \left(0, \sqrt{\frac{m_2 + \dots + m_n}{m_2(m_3 + \dots + m_n)}}, -\sqrt{\frac{m_4 + \dots + m_n}{m_3(m_3 + \dots + m_n)}}, 0, \dots \right) \\ n \geq 5 : \frac{1}{\sqrt{\mu_{34}}} \mathbf{u}_{34} = & \left(0, 0, \sqrt{\frac{m_3 + \dots + m_n}{m_3(m_4 + \dots + m_n)}}, -\sqrt{\frac{m_5 + \dots + m_n}{m_4(m_4 + \dots + m_n)}}, 0, \dots \right) \end{aligned}$$

5.3 Weighted root systems and their metric invariants

It follows from (143)-(147) that the (normalized) Ψ_0 -roots \mathbf{u}_{ik} and \mathbf{u}_{jl} are mutually perpendicular except when $\{i, k\} \cap \{j, l\} \neq \emptyset$, in which case we define $\alpha_k^{i,j} = \alpha_k^{j,i}$ to be the angle between \mathbf{u}_{ik} and \mathbf{u}_{kj} , namely

$$\cos \alpha_k^{i,j} = \mathbf{u}_{ik} \cdot \mathbf{u}_{kj} = -\sqrt{\frac{m_i m_j}{(m_i + m_k)(m_k + m_j)}} \quad (148)$$

In particular, for equal masses the possible angles between any two non-collinear vectors are $\alpha = \frac{1}{2}\pi, \frac{1}{3}\pi$ and $\frac{2}{3}\pi$, and the Ψ_0 -roots constitute, indeed, a root system of type A_{n-1} in the usual sense. For general masses we make the following definition.

Definition 5.4 A weighted root system (of type A_{n-1}) with mass distribution (m_1, m_2, \dots, m_n) is a collection of nonzero vectors

$$\mathbf{w}_{ij} = -\mathbf{w}_{ji}, 1 \leq i < j \leq n,$$

in \mathbb{R}^{n-1} such that i) $\sqrt{\mu_{ij}}|\mathbf{w}_{ij}| = 1$, ii) \mathbf{w}_{ik} and \mathbf{w}_{jl} are perpendicular iff $\{i, k\} \cap \{j, l\} = \emptyset$, and iii) the angle between \mathbf{w}_{ik} and \mathbf{w}_{kj} is $\alpha_k^{i,j}$, as defined by (148).

The Ψ_0 -root system is our prototype of such a root system, and clearly the subset of $n-1$ vectors $\left\{ \frac{1}{\sqrt{\mu_{i,i+1}}} \mathbf{u}_{i,i+1} \right\}$ constitutes a system of *simple roots* for obvious reasons, see (143), (145).

Theorem 5.5 *For a given mass distribution, each Ψ -root system (see Definition 5.1) is a weighted root system in the sense of Definition 5.4, and conversely, each weighted root system is the Ψ -root system for a unique Jacobi transformation $\Psi : M_n \rightarrow M(3, n-1)$. Moreover, for the given mass distribution all the weighted root systems are orthogonally equivalent.*

This result is a simple consequence of the fact that any Ψ can be written uniquely as a composition

$$\Psi = \varphi \circ \Psi_0 : M_n \rightarrow M(3, n-1) \rightarrow M(3, n-1)$$

where $\varphi \in O(n-1)$ acts on $M(3, n-1)$ by multiplication, that is, $\varphi(X) = X\varphi$. Therefore the Jacobi vector matrices X and X' of Ψ_0 and Ψ , respectively, are related by $X' = X\varphi$. Let $\{\pm \mathbf{u}_{ij}\}$ and $\{\pm \mathbf{u}'_{ij}\}$ be the normalized Ψ_0 -roots and Ψ -roots, respectively. The defining relation (142)

$$\sqrt{\mu_{ij}}(\mathbf{a}_i - \mathbf{a}_j) = [\mathbf{u}_{ij}, X] = [\mathbf{u}'_{ij}, X']$$

implies $\mathbf{u}'_{ij} = \mathbf{u}_{ij}\varphi$, and consequently the Ψ -root system is the orthogonally transformed image

$$\left\{ \pm \frac{1}{\sqrt{\mu_{ij}}} \mathbf{u}'_{ij} \right\} = \left\{ \pm \frac{1}{\sqrt{\mu_{ij}}} \mathbf{u}_{ij} \right\} \varphi \quad (\text{row vectors})$$

of the Ψ_0 -root system.

On the other hand, a root system is a finite subset of \mathbb{R}^{n-1} with specified inner products of the vectors, and according to a result of Weyl (see Section 3.3) these numbers determine the subset modulo orthogonal equivalence. Moreover, since the weighted root system spans all \mathbb{R}^{n-1} , any two of them are related by a unique $\varphi \in O(n-1)$. In particular, a weighted root system is the Ψ -root system for a unique Jacobi transformation $\Psi = \varphi \circ \Psi_0$.

5.4 On the role of the mass distribution

The mass distribution of a n -body system manifests itself through the various mass dependent quantities constructed above, for example :

- the collection $\{\mu_{ij}\}$ of reduced masses (29);
- the weighted and normalized root system $\{\pm \mathbf{u}_{ij}\}$;

- the collection $\{\alpha_k^{i,j}\}$ of angles between the vectors of a weighted root system.

The mass distribution (modulo scaling) may, in fact, be reconstructed from any of these invariants. The reduced masses, for example, satisfy the following two conditions

- (i) $\mu_{ij}^{-1} + \mu_{jk}^{-1} > \mu_{ik}^{-1}$ for i, j, k different
- (ii) $\mu_{ij}^{-1} + \mu_{kl}^{-1} = \mu_{ik}^{-1} + \mu_{jl}^{-1} = \mu_{il}^{-1} + \mu_{jk}^{-1}$ for i, j, k, l different

and one can define a *reduced mass distribution of order n* to be a collection of $\binom{n}{2}$ positive numbers $\mu_{ij} = \mu_{ji}$, $1 \leq i < j \leq n$, constrained by the above two conditions. Then there is a 1-1 correspondence between the usual and the reduced mass distributions (modulo scaling) given by the formulae for μ_{ij} and their inversion

$$m_i = \frac{2}{\mu_{ij}^{-1} + \mu_{ik}^{-1} - \mu_{jk}^{-1}} \quad \text{for } i, j, k \text{ different}$$

On the other hand, the congruence space \bar{M}_n (and hence the shape space M_n^*), with the kinematic Riemannian structure, is independent of the mass distribution. The mass distribution determines, however, the representation of congruence classes as points in \bar{M}_n ; in particular, it determines the relative position of the binary collision varieties. Conversely, we can reconstruct the masses m_i (modulo scaling) from knowledge of the position of these varieties.

To illustrate the last statement, consider the case $n = 3$, where M_3^* is a round hemisphere (or sphere, see Section 3.4.4) with a distinguished *equator circle* S^1 representing the shapes of the degenerate triangles. On this circle lie the three collision points b_{12}, b_{23}, b_{13} , where b_{ij} represents the shape of those triangles $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ with $\mathbf{a}_i = \mathbf{a}_j \neq 0$ and the center of mass at the origin. The mass distribution determines their relative position, that is, the angles (or distances) between the points b_{ij} , and once their position have been fixed we know how to determine the position of any shape (cf. [12]).

To describe the relative positions of the points b_{ij} , consider S^1 as the circumscribed circle of the triangle $\Delta = \Delta(b_{12}, b_{23}, b_{31})$ in a Euclidean plane with origin at the center of S^1 . It turns out that Δ is a *central* triangle, in the sense that the center of its circumscribed circle lies in its interior. Moreover, the three central angles

$$0 < \beta_k < \pi, \quad \sum \beta_k = 2\pi$$

where β_k is opposite to the vertex b_{ij} (i, j, k different), are given in terms of the normalized mass distribution by

$$\sin \beta_k = 2 \frac{\sqrt{m_1 m_2 m_3}}{(m_k + m_i)(m_k + m_j)} \quad (\sum m_i = 1)$$

Conversely, any central triangle can be realized in this way and the inversion formula is

$$m_k = 1 - \frac{2 \sin \beta_k}{\sum_{i=1}^3 \sin \beta_i} \quad (\text{cf. [10] or [12]})$$

where the positivity of m_k , indeed, reflects the central property of the triangle Δ .

The triple of angles $(\beta_1, \beta_2, \beta_3)$ is in 1-1 correspondence with another triple $(\alpha_1, \alpha_2, \alpha_3)$, where

$$\alpha_k = \pi - \beta_k/2, \quad k = 1, 2, 3 \quad (149)$$

and these are the central angles of another central triangle, namely the triangle representing (up to congruence) the weighted root system of the mass distribution (cf. Section 5.3). More precisely, whereas $\beta_k = \beta_{ij}$ is the angular distance between b_{ik} and b_{kj} on the circle S^1 , $\alpha_k = \alpha_{ij}$ is the angle between the normalized root vectors \mathbf{u}_{ik} and \mathbf{u}_{kj} since

$$\cos \alpha_k = \mathbf{u}_{ik} \cdot \mathbf{u}_{kj} = -\sqrt{\frac{m_i m_j}{(m_i + m_k)(m_k + m_j)}}, \quad \text{cf. (148)}$$

Finally, we recall that the root system of a k -body system is a weighted root system of type A_{k-1} . The way different A_2 systems combine into higher rank systems A_k , $k \geq 3$, is completely parallel to the "standard" theory of root systems of type A_k , and therefore the higher rank case poses no further problem.

6 Collinear central configurations revisited

Consider the classical Newtonian potential function

$$U = \sum_{i < j} \frac{m_i m_j}{|\mathbf{a}_i - \mathbf{a}_j|} \quad (150)$$

and its gradient field ∇U with respect to the kinematic metric (8) in M_n (or \hat{M}_n). An n -configuration $\mathbf{X} = (\mathbf{a}_1, \dots, \mathbf{a}_n)$ is called *central* if $\nabla U(\mathbf{X}) = \lambda \mathbf{X}$ for some constant λ , namely

$$\lambda \mathbf{a}_i = \frac{1}{m_i} \frac{\partial U}{\partial \mathbf{a}_i} = \sum_{j \neq i} \frac{m_j (\mathbf{a}_j - \mathbf{a}_i)}{|\mathbf{a}_i - \mathbf{a}_j|^3}, \quad i = 1, \dots, n \quad (151)$$

In fact, $\lambda = -I^{-1}U(\mathbf{X})$ where $I = |\mathbf{X}|^2$, since by a classical result of Euler the homogeneity of U implies $\nabla U(\mathbf{X}) \cdot \mathbf{X} = -U(\mathbf{X})$. The collection of central configurations is clearly invariant under similarity transformations. Thus we may fix a scaling of the vectors, say $I = 1$, and ask for solutions modulo $O(3)$ -congruence. Then the solutions are just the critical points of U as a function restricted to the shape space M_n^* . For $n = 3$ the only solution which is the shape of a non-degenerate triangle is the equilateral triangle (by Lagrange, 1772), whereas there are three degenerate triangle solutions (by Euler, 1767) and they are represented by the three *Euler points* on the equator circle of M_3^* .

Even the enumeration of all critical points in M_n^* is, in fact, still an open problem for $n \geq 4$, and now the number also depends on the mass distribution. It is only known to be finite for $n < 5$. However, the number of collinear solutions is known to be $n!/2$ for all $n \geq 3$, and the first proof was presented

by Moulton [23]. More recently, Smale [26] has given a topological proof using elementary Morse theory. In this subsection we shall give a similar and quite simple proof using the weighted root system of an n-body system.

As usual, we assume that U in (150) is a function of n vector variables \mathbf{a}_i linearly related by the condition that the center of mass is at the origin. Actually this condition is a consequence of the identities (151). Anyhow, let us express U as a function of the $n-1$ Jacobi vectors

$$X = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1}) = [\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_3^*]$$

defined by a fixed Jacobi transformation

$$\Psi : M_n \rightarrow M(3, n-1)$$

with associated root system $\{\pm \mathbf{w}_{ij}\}$. Take, for example, the standard transformation Ψ_0 and the expressions for the vectors \mathbf{x}_i and \mathbf{w}_{ij} in Section 2.3 and Section 5.2. The (row) vectors \mathbf{w}_{ij} and \mathbf{x}_k^* belong to the same Euclidean space \mathbb{R}^{n-1} and by definition

$$|\mathbf{a}_i - \mathbf{a}_j|^2 = \sum_{k=1}^3 (\mathbf{w}_{ij} \cdot \mathbf{x}_k^*)^2, \quad \mathbf{w}_{ij} = \frac{1}{\sqrt{\mu_{ij}}} \mathbf{u}_{ij} \quad (\text{cf. Section 5.1})$$

The gradient field ∇U is tangential to the subvariety of collinear configurations and therefore the critical points we seek are also the critical points of U restricted to the subvariety of collinear shapes, namely $M^*(1, n-1) \simeq RP^{n-2}$. Let us represent all shapes of collinear type by n -configurations $\mathbf{X} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ in M_n with position vectors \mathbf{a}_i along the x -axis, and therefore the Jacobi matrix X has row vectors $\mathbf{x}_2^* = \mathbf{x}_3^* = 0$. Hence, we shall regard U as a function on the $(n-1)$ -space

$$\mathbb{R}^{n-1} : \mathbf{x} = \mathbf{x}_1^* = (x_1, \dots, x_{n-1})$$

and the condition $I = 1$ means restriction to the unit sphere $S^{n-2} : \sum x_i^2 = 1$. When antipodal points on this sphere are identified, we obtain the above projective space RP^{n-2} , see (154) below.

For convenience, we shall identify the root vector \mathbf{w}_{ij} with the linear functional

$$\omega_{ij} : \mathbb{R}^{n-1} \rightarrow \mathbb{R}, \quad \omega_{ij}(\mathbf{x}) = \mathbf{w}_{ij} \cdot \mathbf{x},$$

and regard the family

$$\Delta = \{\pm \omega_{ij}\} = \Delta_+ \cup \Delta_-$$

as the *root system*. The *positive* and *simple* roots are the subfamilies

$$\Delta_+ = \{\omega_{ij}, i < j\} \supset \Pi = \{\omega_{12}, \omega_{23}, \dots, \omega_{n-1,n}\},$$

respectively, with $\omega_{ij} = -\omega_{ji}$ and $\omega_{ij} + \omega_{jk} = \omega_{ik}$, and following the "usual" procedure we define hyperplanes

$$H_{ij} = \ker \omega_{ij} \subset \mathbb{R}^{n-1}$$

which subdivide the space into disjoint, open and connected chambers C_i whose union is

$$C_1 \cup C_2 \cup \dots = \mathbb{R}^{n-1} - \cup_{i < j} H_{ij} .$$

Clearly, there are $n!$ chambers, and each chamber C is distinguished by a specific choice of signs $\varepsilon_{ij} = \pm 1$ such that

$$\mathbf{x} \in C \Leftrightarrow \varepsilon_{ij} \omega_{ij}(\mathbf{x}) > 0 \text{ for all } i < j$$

By changing all signs one obtains the antipodal chamber. In particular, our *fundamental chamber* C_0 is the one with all ε_{ij} positive, namely

$$C_0 : \omega_{ij}(\mathbf{x}) > 0 \text{ for all } i < j,$$

or equivalently in terms of the positions $\mathbf{a}_i = a_i$ of the point masses (on the x-axis), since $\omega_{ij}(\mathbf{x}) = a_i - a_j$,

$$a_1 > a_2 > \dots > a_n \quad (152)$$

The above decomposition and combinatorial structure is, of course, similar to the Weyl chamber decomposition for the Weyl group $A_{n-1} = S_n$ of $SU(n)$.

By (140) and (141), the expression (150) may be written

$$U(\mathbf{x}) = \sum_{i < j} \frac{m_i m_j}{|\omega_{ij}(\mathbf{x})|} \quad (153)$$

which in the fundamental chamber reads

$$U(\mathbf{x}) = \sum_{i < j} \frac{m_i m_j}{\omega_{ij}(\mathbf{x})} \text{ for } \mathbf{x} \in C_0$$

Lemma 6.1 *$U(\mathbf{x})$ is a convex function on the chamber C_0 , and it has a unique critical point \mathbf{x}_0 (where U has a minimum) in the spherical chamber $C_0 \cap S^{n-2}$.*

Proof. By straightforward calculations

$$\begin{aligned} \frac{\partial U}{\partial x_k} &= - \sum_{i < j} m_i m_j \frac{w_{ij}^k}{\omega_{ij}(\mathbf{x})^2} \\ \frac{\partial^2 U}{\partial x_k \partial x_l} &= 2 \sum_{i < j} m_i m_j \frac{w_{ij}^k w_{ij}^l}{\omega_{ij}(\mathbf{x})^3} \end{aligned}$$

where $\mathbf{w}_{ij} = (w_{ij}^1, w_{ij}^2, \dots, w_{ij}^{n-1})$ and $\mathbf{x} = (x_1, x_2, \dots, x_{n-1})$. Hence, the Hessian of U at $\mathbf{x} \in C_0$ is the following positive definite quadratic form in the variable $\mathbf{t} = (t_1, t_2, \dots, t_{n-1})$

$$HU(\mathbf{x})(\mathbf{t}) = \frac{1}{2} \sum_{k,l=1}^{n-1} \frac{\partial^2 U}{\partial x_k \partial x_l}(\mathbf{x}) t_k t_l = \sum_{i < j} m_i m_j \frac{\omega_{ij}(\mathbf{t})^2}{\omega_{ij}(\mathbf{x})^3}$$

On the other hand, $U(\mathbf{x}) > 0$ on C_0 and $U(\mathbf{x}) \rightarrow \infty$ as \mathbf{x} approaches the walls of the chamber and is bounded away from the origin. It follows that U

has a unique critical point in the spherical chamber, namely a minimumspoint \mathbf{x}_0 . ■

The induced congruence group action on \mathbb{R}^{n-1} is simply the group $O(1) = \{\pm 1\}$ acting by inversion $\mathbf{x} \rightarrow -\mathbf{x}$, and therefore antipodal points on the sphere S^{n-2} represent collinear n-configurations with the same shape :

$$S^{n-2} \rightarrow S^{n-2}/O(1) = \mathbb{R}P^{n-2} \subset M_n^* \quad (154)$$

In particular, a pairs of antipodal chambers is mapped to the same chamber in $\mathbb{R}P^{n-2}$, and consequently the latter space is divided into $n!/2$ chambers, each containing a unique critical point of U as a function on $\mathbb{R}P^{n-2}$. This completes the proof of the enumeration result originally due to Moulton.

Remark 6.2 *The above central configuration solutions (a_1, a_2, \dots, a_n) can be distinguished by the ordering $a_{i_1} > a_{i_2} > \dots > a_{i_n}$ (modulo inversion of order), and the minimumspoint \mathbf{x}_0 in the fundamental chamber C_0 is the solution with the ordering (152). However, for a given mass distribution (m_1, m_2, \dots, m_n) the solution satisfying $a_{i_1} > a_{i_2} > \dots > a_{i_n}$ can also be found by the same procedure, namely as the minimumspoint \mathbf{x}_0 in the fundamental chamber C_0 corresponding to the permuted mass distribution $(m_{i_1}, m_{i_2}, \dots, m_{i_n})$. In particular, for each string of equal masses $m_i = m_j = \dots = m_k$ the set of solutions is invariant under permutations of a_i, a_j, \dots, a_k .*

Let us briefly consider the explicit numerical calculation of the collinear central configurations by calculating the critical points of U on S^{n-2} . Observe that the identity (151) with the additional condition $I = 1$ may be interpreted as an application of the classical Lagrange multiplier method with a constraint. Similarly, in the present case where we restrict to collinear configurations, we seek the $n!$ solutions of the system

$$\nabla U(\mathbf{x}) = \lambda \mathbf{x}, \quad |\mathbf{x}| = 1 \quad (155)$$

where $\lambda = -U(\mathbf{x})$ (by Euler's formula) and

$$\nabla U(\mathbf{x}) = - \sum_{i < j} m_i m_j \frac{\mathbf{w}_{ij}}{\omega_{ij}(\mathbf{x})^2}$$

The condition for a solution is the vanishing of the component of the gradient tangential to the sphere, namely

$$\nabla U(\mathbf{x}) + U(\mathbf{x})\mathbf{x} = 0 \quad (156)$$

As a simpleminded algorithm for finding a solution we construct a sequence of points $\mathbf{y}_k \in S^{n-2}$ by "moving" on the sphere in the direction opposite to the vector (156), as follows

$$\mathbf{y}_{k+1} = \frac{(1 - U(\mathbf{y}_k))\mathbf{y}_k - \nabla U(\mathbf{y}_k)}{|(1 - U(\mathbf{y}_k))\mathbf{y}_k - \nabla U(\mathbf{y}_k)|},$$

starting from an initial point \mathbf{y}_1 in a given chamber. In general, the sequence will stay in the chamber and converge to the minimum point of U . The algorithm can certainly be made more effective, but we shall not discuss these matters.

Another approach is to introduce new variables t_i linearly related to x_1, x_2, \dots, x_{n-1} via the simple roots.

$$t_i = \omega_{i,i+1}(\mathbf{x}) = a_i - a_{i+1}, \quad i = 1, 2, \dots, n-1$$

For example, in the case $n = 3$ the system (155) reads

$$\begin{aligned} 0 &= -\frac{m_1}{(t_1)^2} + \frac{1-m_1}{(t_2)^2} + \frac{m_1}{(t_1+t_2)^2} - \lambda t_2 \\ 0 &= \frac{m_2}{(t_1)^2} + \frac{m_3}{(t_1+t_2)^2} - \lambda[(1-m_1)t_1 + m_3 t_2] \end{aligned} \quad (157)$$

as compared to the initial system (151) which has three dependent equations similar to the above ones. In fact, the first equation of system (151) and (157) are identical. By elimination of λ we obtain a homogeneous equation of degree 5 which has a unique solution (modulo scaling) with both $t_i > 0$. Thus it can be reduced to an inhomogeneous equation of degree 4, whereas Euler's approach led to the solution of a 5th order inhomogeneous equation $P(\omega) = 0$ (cf.[25], §14). The latter equation is, in fact, obtained from the system (157) when we write $(t_1 + t_2) = t$, $t_1 = \omega t$, $t_2 = (1 - \omega)t$ and eliminate the resulting multiplier $\lambda' = \lambda t^3$.

References

- [1] M. Atiyah and J. Berndt, *Projective planes, Severi varieties and spheres*, arXiv: math.DG/0206135; in *Surveys in Differential Geometry*, vol. VIII, International Press 2003.
- [2] V.I. Arnold, *Ramified Covering $\mathbb{CP}^2 \rightarrow S^4$, Hyperbolicity and Projective Topology*, Sib. Mat. Zh., 1988, vol. 29, no. 5, pp. 36-47.
- [3] V. I. Arnold, *Relatives of the Quotient of the Complex Projective Plane by the Complex Conjugation*, Proc. Steklov Inst. Math. **224** (1999), 46-56.
- [4] A. Back and W. Y. Hsiang, *Equivariant geometry and Kervaire spheres*, Transactions AMS **304** (1987), 207-227.
- [5] B. H. Bransden and C. J. Joachain, *Physics of atoms and molecules*, Longman, London and New York 1983.
- [6] B. H. Bransden and C. J. Joachain, *Introduction to Quantum Mechanics*, Longman Scientific & Technical, 1994.
- [7] G. Bredon, *Introduction to compact transformation groups*, Academic Press, 1972.
- [8] W. Y. Hsiang, *Geometric study of the three-body problem*, Center for Pure and Applied Mathematics, Univ. of Calif., Berkeley, PAM- 620, 1994.

- [9] W. Y. Hsiang, *On the kinematic geometry of many body systems*, PAM-719 preprint at U.C. Berkeley, Nov. 1997; Chinese Ann. Math. Ser. B **20** (1999), no. 1, 11-28.
- [10] W. Y. Hsiang and E. Straume, *Kinematic geometry of triangles with given mass distribution*, PAM-636 preprint at U. C. Berkeley, May 1995.
- [11] W. Y. Hsiang and E. Straume, *Topology and geometry of many body systems*, preprint at Dept. of Math., NTNU, June 2000.
- [12] W. Y. Hsiang and E. Straume, *Kinematic geometry of triangles and the study of the three-body problem*, Dept. of Math., NTNU, June 2004; arXiv: math-ph/0608060.
- [13] T. Iwai, *A geometric setting for classical molecular dynamics*, Ann. Inst. Henry Poincaré, **47** (1987), 199-219.
- [14] T. Iwai, *A geometric setting for the quantum planar n -body system, and a $U(n-1)$ basis for the internal states*, J. Math. Phys. **29**(6), June 1988.
- [15] N. H. Kuiper, *The Quotient Space of $\mathbb{CP}(2)$ by Complex Conjugation is the 4-Sphere*, Math. Ann. **208** (1974), 175-177.
- [16] R. G. Littlejohn and M. Reinsch, *Internal or shape coordinates in the n -body problem*, Phys. Rev., A **52** (1995), 2035-2051.
- [17] R. G. Littlejohn and M. Reinsch, *Gauge fields in the separation of rotations and internal motions in the n -body problem*, Reviews of Modern Physics **69** (1997), no.1, 213-275.
- [18] K. A. Mitchell and R. G. Littlejohn, *Kinematic orbits and the structure of the internal space for systems of five or more bodies*, J. Phys. A: Math. Gen. **33** (2000) 1395-1416.
- [19] J. Lützen, *Interactions between mechanics and differential geometry in the 19th century*, Arch. Hist. Exact Sci. **49** (1995), no. 1, 1-72.
- [20] C. Marchal, *The Three-Body Problem*, Studies in Astronautics 4, Elsevier, Amsterdam, 1990.
- [21] W. S. Massey, *The quotient space of the complex projective plane under conjugation is a 4-sphere*, Geometriae Dedicata **2** (1973), 371-374.
- [22] J. Milnor, *Lectures on the h -cobordism theorem*, Mathematical Notes, Princeton Univ. Press, 1965.
- [23] F. R. Moulton, *The straight line solutions of the problem of n bodies*, Annals of Math., **12**, 1-17 (1910).
- [24] G. Schwarz, *Smooth functions invariant under the action of a compact Lie group*, Topology **14** (1975), 63-68.

- [25] C.L. Siegel and J.K. Moser, *Lectures on Celestial Mechanics*, Springer-Verlag, Berlin 1971.
- [26] S. Smale, *Topology of Mechanics II*; Inventiones math. **11**, 45-64 (1970).
- [27] E. Straume, *Weyl groups and the regularity properties of certain compact Lie group actions*, Transactions AMS **306** (1988), 165 -190.
- [28] E. Straume, *On the invariant theory and geometry of compact linear groups of cohomogeneity ≤ 3* , Differential Geometry and its Applications **4** (1994), 1-23.
- [29] E. Straume, *Compact Connected Lie Transformation Groups on Spheres with Low Cohomogeneity, I*, Memoirs of AMS, Vol.119, No.569, 1996; *II*, Memoirs of AMS, Vol.125, No.595, 1997.
- [30] E. Straume, *On the geometry and behavior of n-body motions*, International Journal of Mathematics and Mathematical Sciences **28:12** (2001) 689-732.
- [31] H. Weyl, *The Classical Groups*, second edition, Princeton University Press, Princeton 1946.